

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)

77515

Delaval, Jan

From: Roark, Jessica
Sent: Wednesday, October 09, 2002 8:49 AM
To: Delaval, Jan
Subject: 09/806955

Jan,

Please search, including pending, the following from 09/806955:

SEQ ID NO:1 and
SEQ ID NO:2.

Results on paper please.

Thanks!

Jessica H. Roark

CM1 8A03
Mailbox 9E12
Art Unit 1644
703 605-1209

SEARCH
SEQ 2 = SEQ 1 except
SEQ 1 has a C-atom 6 His tag!

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:10:24 ; Search time 26.02 Seconds
(without alignments)
599,845 Million cell updates/sec

Title: US-09-806-955A-1

Perfect score: 3273
Sequence: 1 MEEDKKEVDGVGVGIDIGTT.....AGPPTGEEDTAELHHHHH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

1 number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing First 45 summaries

Database : Issued Patents,AA:*

1: /cgn2_6/prodata/2/1aa/3A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCUTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	98.0	654	1	US-08-441-139-11 Sequence 11, Appl
2	3168	96.8	666	1	US-08-441-139-16 Sequence 16, Appl
3	2174	66.4	682	1	US-08-441-139-2 Sequence 2, Appl
4	2156	65.9	663	1	US-08-441-139-7 Sequence 7, Appl
5	2077	63.5	646	1	US-08-441-139-14 Sequence 14, Appl
6	2065	63.1	679	1	US-08-441-139-5 Sequence 5, Appl
7	1938	59.2	643	4	US-08-797-358B-3 Sequence 3, Appl
8	1547.5	47.3	679	1	US-08-214-583-2 Sequence 2, Appl
9	1474	45.0	641	1	US-08-441-139-4 Sequence 4, Appl
10	1461.5	44.7	649	4	US-09-066-047-5 Sequence 5, Appl
11	1444	44.1	607	2	US-08-472-534-5 Sequence 5, Appl
12	1303.5	39.8	600	6	5240706-1 Patent No. 5240706
13	1140.5	34.8	339	2	US-08-928-692-52 Sequence 52, Appl
14	941	28.8	187	6	5196523-13 Patent No. 5196523
15	824	25.2	168	1	US-08-441-139-10 Sequence 10, Appl
16	818.5	25.0	315	1	US-08-257-073-7 Sequence 7, Appl
17	750.5	22.9	471	1	US-08-303-905B-2 Sequence 2, Appl
18	726.5	22.2	472	1	US-08-203-905B-14 Sequence 14, Appl
19	701	21.4	307	4	US-08-858-207A-481 Patent No. 481, App
20	642.5	19.6	129	6	5196523-10 Patent No. 5196523
21	607.5	18.6	999	3	US-08-770-301A-3 Sequence 3, Appl
22	607.5	18.6	999	3	US-09-175-581-3 Sequence 3, Appl
23	598	18.3	999	3	US-08-770-301A-1 Sequence 1, Appl
24	598	18.3	999	3	US-09-175-581-1 Sequence 1, Appl
25	528.5	16.1	560	2	US-08-928-692-53 Sequence 53, Appl
26	437.5	13.4	374	2	US-08-928-692-51 Sequence 51, Appl
27	402	12.3	77	6	5196523-7 Patent No. 5196523

28	395	12.1	79	6	5196523-11 Patent No. 5196523
29	367.5	11.2	599	2	US-09-080-983-9 Sequence 9, Appl
30	358	10.9	549	2	US-08-770-544-6 Sequence 6, Appl
31	343.5	10.6	80	1	US-08-464-164-4 Sequence 4, Appl
32	345.5	10.6	80	1	US-08-338-057-4 Sequence 4, Appl
33	345.5	10.6	80	2	US-08-668-416-4 Sequence 4, Appl
34	323	9.9	554	5	PCT-US94-06430-7 Sequence 7, Appl
35	320	9.8	554	4	US-08-591-668-7 Sequence 7, Appl
36	267.5	8.2	714	2	US-08-472-534-3 Sequence 3, Appl
37	264	8.1	56	6	5196523-12 Patent No. 5196523
38	230	7.0	46	6	5196523-8 Patent No. 5196523
39	211	6.4	42	6	5196523-15 Patent No. 5196523
40	192	5.9	38	6	5196523-9 Patent No. 5196523
41	168.5	5.1	1786	4	US-08-973-462-8 Sequence 8, Appl
42	166	5.1	125	4	US-08-858-207A-407 Sequence 407, App
43	151.5	4.6	1805	1	US-07-853-913-2 Sequence 2, Appl
44	147	4.5	693	4	US-08-235-836C-68 Sequence 68, Appl
45	142.5	4.4	2285	4	US-09-308-375-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-441-139-11
Sequence 11, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELEPHONE: 516-742-4343
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-11
Query Match 98.0%; Score 3209; DB 1; Length 654;
Best Local Similarity 99.7%; Pred. No. 5,2e-250;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Caps 0;

QY 2 EEDKEDVGTGVGIDLTGTYSCVGFKNGRVEIIANDOGNRITPSYVAFTEPGERLIGDA 61
DB 20 EEDKEDVGTGVGIDLTGTYSCVGFKNGRVEIIANDOGNRITPSYVAFTEPGERLIGDA 79
QY 62 AKNOLTSNPENTVPDAKRLIGRTWNPVSVOODIKFLPKRVVEKTKPKYIOVDIGGQRT 121
DB 80 AKNOLTSNPENTVPDAKRLIGRTWNPVSVOODIKFLPKRVVEKTKPKYIOVDIGGQRT 139
QY 122 FAPETISAMVLTAKMETEAVLGRKVTAAVTVPAVFNDARQAKDAGTAGLVNMRIT 181
DB 140 FAPETISAMVLTAKMETEAVLGRKVTAAVTVPAVFNDARQAKDAGTAGLVNMRIT 199
QY 182 NEPTAAAIAYGLDKRREGKNILVFDLGGSTFDVSLTTIDNGFEVYATNGDTHLGGEDPD 241
DB 200 NEPTAAAIAYGLDKRREGKNILVFDLGGSTFDVSLTTIDNGFEVYATNGDTHLGGEDPD 259
QY 242 QRVMEHFILKLYKKTKGKDVKRDNRVAVOKLRREVEKAKALSSOHQARIETESFEGEDPS 301
DB 260 QRVMEHFILKLYKKTKGKDVKRDNRVAVOKLRREVEKAKALSSOHQARIETESFEGEDPS 319
QY 302 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSTRIPKIQOLVKEFF 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSTRIPKIQOLVKEFF 379
QY 362 NGKEPSRCINPDEAVAYGAAGVAGVLSGDDPTGDLVLLDVCPLTGLIETVGVMTKLIPR 421
DB 380 NGKEPSRCINPDEAVAYGAAGVAGVLSGDDPTGDLVLLDVCPLTGLIETVGVMTKLIPR 439
QY 422 NTVPVTKKSQIFSTASDNPVTITIKVYEGEPLTNDNHLGTFDLTGIPPARGVPOIEV 481
DB 440 NTVPVTKKSQIFSTASDNPVTITIKVYEGEPLTNDNHLGTFDLTGIPPARGVPOIEV 499
QY 482 TFEIDVNGILRYTAEDKGTGNKKTITINDONRLPPEIERVAVNDAEKFAEDDKLKIKERI 541
DB 500 TFEIDVNGILRYTAEDKGTGNKKTITINDONRLPPEIERVAVNDAEKFAEDDKLKIKERI 559
QY 542 DTRNELESYASLKNQIDGKLGKLSSEDEKTEKAVEKEIEMLESHODADIEDFAK 601
DB 560 DTRNELESYASLKNQIDGKLGKLSSEDEKTEKAVEKEIEMLESHODADIEDFAK 619
QY 602 KKELEIVQPIISKLXGSAAGPPTEGDETA 632
DB 620 KKELEIVQPIISKLXGSAAGPPTEGDETA 650

RESULT 2
US-08-441-139-16
Sequence 16, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: WITTUP, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441.139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-16

Query Match 96.8%; Score 3168; DB 1; Length 666;
Best Local Similarity 97.8%; Pred. No. 1,1e-246;
Matches 617; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

QY 2 EEDKEDVGTGVGIDLTGTYSCVGFKNGRVEIIANDOGNRITPSYVAFTEPGERLIGDA 61
DB 32 DEEDKEDVGTGVGIDLTGTYSCVGFKNGRVEIIANDOGNRITPSYVAFTEPGERLIGDA 91
QY 62 AKNOLTSNPENTVPDAKRLIGRTWNPVSVOODIKFLPKRVVEKTKPKYIOVDIGGQRT 121
DB 92 AKNOLTSNPENTVPDAKRLIGRTWNPVSVOODIKFLPKRVVEKTKPKYIOVDIGGQRT 151
QY 122 FAPETISAMVLTAKMETEAVLGRKVTAAVTVPAVFNDARQAKDAGTAGLVNMRIT 181
DB 152 FAPETISAMVLTAKMETEAVLGRKVTAAVTVPAVFNDARQAKDAGTAGLVNMRIT 211
QY 182 NEPTAAAIAYGLDKRREGKNILVFDLGGSTFDVSLTTIDNGFEVYATNGDTHLGGEDPD 241
DB 212 NEPTAAAIAYGLDKRREGKNILVFDLGGSTFDVSLTTIDNGFEVYATNGDTHLGGEDPD 271
QY 242 QRVMEHFILKLYKKTKGKDVKRDNRVAVOKLRREVEKAKALSSOHQARIETESFEGEDPS 301
DB 272 QRVMEHFILKLYKKTKGKDVKRDNRVAVOKLRREVEKAKALSSOHQARIETESFEGEDPS 331
QY 302 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSTRIPKIQOLVKEFF 361
DB 332 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSTRIPKIQOLVKEFF 391
QY 362 NGKEPSRCINPDEAVAYGAAGVAGVLSGDDPTGDLVLLDVCPLTGLIETVGVMTKLIPR 421
DB 392 NGKEPSRCINPDEAVAYGAAGVAGVLSGDDPTGDLVLLDVCPLTGLIETVGVMTKLIPR 451
QY 422 NTVPVTKKSQIFSTASDNPVTITIKVYEGEPLTNDNHLGTFDLTGIPPARGVPOIEV 481
DB 452 NTVPVTKKSQIFSTASDNPVTITIKVYEGEPLTNDNHLGTFDLTGIPPARGVPOIEV 511
QY 482 TFEIDVNGILRYTAEDKGTGNKKTITINDONRLPPEIERVAVNDAEKFAEDDKLKIKERI 541
DB 512 TFEIDVNGILRYTAEDKGTGNKKTITINDONRLPPEIERVAVNDAEKFAEDDKLKIKERI 571
QY 542 DTRNELESYASLKNQIDGKLGKLSSEDEKTEKAVEKEIEMLESHODADIEDFAK 601
DB 572 DTRNELESYASLKNQIDGKLGKLSSEDEKTEKAVEKEIEMLESHODADIEDFAK 631
QY 602 KKELEIVQPIISKLXGSAAGPPTEGDETA 632
DB 632 KKELEIVQPIISKLXGSAAGPPTEGDETA 662

RESULT 3
US-08-441-139-2
Sequence 2, Application US/08441139
Patent No. 5773245

[illegible]

```

1      RESULT 4
2      US-08-441-139-7
3      Sequence 7, Application US/08441139
4      Patent No. 5773245
5      GENERAL INFORMATION:
6      APPLICANT: Wiltup, Dr. Karl D.
7      APPLICANT: Robinson, Anne S.
8      TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
9      TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
10     NUMBER OF SEQUENCES: 20
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
13     STREET: 400 Garden City Plaza
14     CITY: Garden City
15     STATE: NY
16     COUNTRY: USA
17     ZIP: 11530
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: PatentIn Release #1.0, Version #1.25
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/441,139
25     FILING DATE: 15-MAY-1995
26     CLASSIFICATION: 435
27     PRIOR APPLICATION DATA:
28     APPLICATION NUMBER: US 08/089,997
29     FILING DATE: 06-JUL-1993
30     ATTORNEY/AGENT INFORMATION:
31     NAME: Digilio, Frank S.
32     REGISTRATION NUMBER: 31,346
33     REFERENCE/DOCKET NUMBER: 8646
34     TELECOMMUNICATION INFORMATION:
35     TELEPHONE: 516-742-4343
36     TELEFAX: 516-742-4366
37     TELEX: 230 901 SANS UR
38     INFORMATION FOR SEQ ID NO: 7:
39     SEQUENCE CHARACTERISTICS:
40     LENGTH: 663 amino acids
41     TYPE: amino acid
42     TOPOLOGY: linear
43     MOLECULE TYPE: protein
44     US-08-441-139-7

```

[illegible]

```

Db 86 AANQAPSNENITFDIKRLIGRKFDEKTMADIKSPFPIVNDKNPDLVAVN-GGKKRK 144
Qy 122 FAPPEISAVLTMMKETAAYLGKVTTHAVVTPAYFNDAORATDAGTIGLVNMTI 181
Db 145 FTPEEISAMILSKMOTAAVYLGKPTTHSVVTPAYFNDAORATDAGTIGLVNMTI 204
Qy 182 NEPTAAAIYVGLDREGENILVFDLGGCTPDVSLTTIDNGVEVAVTNGDTHLGDEPD 241
Db 205 NEPTAAAIYVGLDREGENILVFDLGGCTPDVSLTTIDNGVEVAVTNGDTHLGDEPD 264
Qy 242 QVMEHFIKLYKKTKGKDVKNRAVOKLRREYKAKRALSSQHOARIELESEYEEDEFS 301
Db 265 NRVINLTATYNNRKNVDYTKDLKAMGKLKREYKANGTLLSQSVRIEISEFNFQDPS 324
Qy 302 ETLTRAKFEELNMDLFRSTMKPVQKVLKEDSDLKSDIDEIVLGSGSTRIPKIQOLYKEFF 361
Db 325 ETLTRAKFEELNMDLFRSTMKPVQKVLKEDSDLKSDIDEIVLGSGSTRIPKIQOLYKEFF 384
Qy 362 NCKEPRSGINPDEAVAYGAAYVAGVLSGDDGTGDLVLDVCPPLTIGETVGVMTKLPR 421
Db 385 -CKKASKGINPDEAVAYGAAYVAGVLSGEGSDNIVLLDVLPLTIGETVGVMTKLGR 443
Qy 422 NRVVPTKKQIFSTADNCPVTYIKYEGEERPLTKNHLGTFDLTGIPAPRGVQIYEV 481
Db 444 NRVVPTKKQIFSTADNCPVTYIKYEGEERPLTKNHLGTFDLTGIPAPRGVQIYEV 503
Qy 482 TFEIDVNGILRYTAEDK-CTGNMKNITITNDQNRLLTPEIERMVNDAEKFAEDKTLKR 540
Db 504 TFEIDVNGILRYTAEDK-CTGNMKNITITNDQNRLLTPEIERMVNDAEKFAEDKTLKR 563
Qy 541 ITRNELEYAVSLKNOICDKELGKLSSEDEKTEKAVEEKEIEMLESH-QDADIEDPK 599
Db 564 ITRNELEYAVSLKNOICDKELGKLSSEDEKTEKAVEEKEIEMLESH-QDADIEDPK 623
Qy 600 AKKELEEVPIISKLVSAGPPPTGEED 629
Db 624 DQROKIDAVVHPITOKLY-SEGAGDADEED 652

```

RESULT 5

```

US-08-441-139-14
: Sequence 14, Application US/08441139
: Patent No. 5773245
: GENERAL INFORMATION:
: APPLICANT: Wiltup, Dr. Karl D.
: APPLICANT: Robinson, Anne S.
: TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
: TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
: STREET: 400 Garden City Plaza
: CITY: Garden City
: STATE: NY
: COUNTRY: USA
: ZIP: 11530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PC-DOS/MS-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/441.139
: FILING DATE: 15-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/089,997
: FILING DATE: 06-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: DIGILIO, Frank S.
: REGISTRATION NUMBER: 31,346
: REFERENCE/DOCKET NUMBER: 8646

```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 516-742-4343
: TELEFAX: 516-742-4366
: TELEX: 230 901 SANS UR
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 646 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-441-139-14

```

```

Query Match 63.5%; Score 2077; DB 1; Length 646;
Best Local Similarity 66.0%; Pred. No. 6,5e-159;
Matches 409; Conservative 97; Mismatches 108; Indels 6; Gaps 5;

```

```

Qy 10 GIVVIGIDLTGTYSCVGFKNRVEITANDGNRTTPSYAFPEGERLIGDAKNOLTSN 69
Db 4 GNAVIGIDLTGTYSCVGFKNRVEITANDGNRTTPSYAFPEGERLIGDAKNOLTSN 62
Qy 70 PENTVFDAKRLIGRTNDSVQODIKFLPKVYEKTKPIQVYDIDGGCTKTFAPPEISA 129
Db 63 PNTVFDAKRLIGRTNDSVQODIKFLPKVYEKTKPIQVYDIDGGCTKTFAPPEISA 121
Qy 130 MVLTKKETAEAYLGKVTTHAVVTPAYFNDAORATDAGTIGLVNMTI 189
Db 122 MVLTKKETAEAYLGKVTTHAVVTPAYFNDAORATDAGTIGLVNMTI 181
Qy 190 AYGLDKREG-EKILVYFDLGGCTPDVSLTTIDNGVEVAVTNGDTHLGDEPDQRMNEH 248
Db 182 AYGLDKREG-EKILVYFDLGGCTPDVSLTTIDNGVEVAVTNGDTHLGDEPDQRMNEH 241
Qy 249 IKLYKKTKGKDVKNRAVOKLRREYKAKRALSSQHOARIELESEYEEGEFSETLTKR 308
Db 242 IAFKRRKKHKKDISENKRAVRRLTACERAKRLTSSQASIEIDSLXEGIDFTSTIRAR 301
Qy 309 FEELNMDLFRSTMKPVQKVLKEDSDLKSDIDEIVLGSGSTRIPKIQOLYKEFFNGKPSR 368
Db 302 FEELNMDLFRSTMKPVQKVLKEDSDLKSDIDEIVLGSGSTRIPKIQOLYKEFFNGKPSR 361
Qy 369 GINPDEAVAYGAAYVAGVLSGDDGTGDLVLDVCPPLTIGETVGVMTKLPR 426
Db 362 GINPDEAVAYGAAYVAGVLSGDDGTGDLVLDVCPPLTIGETVGVMTKLPR 421
Qy 427 TKKSQIFSTADNCPVTYIKYEGEERPLTKNHLGTFDLTGIPAPRGVQIYEVFEID 486
Db 422 TKKSQIFSTADNCPVTYIKYEGEERPLTKNHLGTFDLTGIPAPRGVQIYEVFEID 481
Qy 487 VNGILRYTAEDKCTGNMKNITITNDQNRLLTPEIERMVNDAEKFAEDKTLKERIDTRNE 546
Db 482 VNGILRYTAEDKCTGNMKNITITNDQNRLLTPEIERMVNDAEKFAEDKTLKERIDTRNE 541
Qy 547 LESYAVSLKNOICDKELGKLSSEDEKTEKAVEEKEIEMLESH-QDADIEDFAKKKELE 606
Db 542 LESYAVSLKNOICDKELGKLSSEDEKTEKAVEEKEIEMLESH-QDADIEDFAKKKELE 600
Qy 607 EIVQPIISKLVSAGPPPTG 626
Db 601 KVCNPIITTKLYSAGGMPG 620

```

RESULT 6

```

US-08-441-139-5
: Sequence 5, Application US/08441139
: Patent No. 5773245
: GENERAL INFORMATION:
: APPLICANT: Wiltup, Dr. Karl D.
: APPLICANT: Robinson, Anne S.
: TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
: TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:

```


ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-5

Query Match 63.1%; Score 2065.5; DB 1; Length 679;
Best Local Similarity 65.4%; Pred. No. 5.9e-158;
Matches 403; Conservative 90; Mismatches 118; Indels 5; Gaps 4;

QY 7 EDVGTGVIDIGTTSYSCVGFNGRVEIIANDGNRIIPSYAFTPEGERLIGDAKNOL 66
DB 48 EDYGVIGIDIGTTSYSCVAVMKNKTEILANDGNRIIPSYFT-DDERLIGDAKNQA 106
QY 67 TSNPNTYFPAKRLIGRTWNPDSVOODIKFLPKYVEKTKRYIOVDIGGQTKTFAPE 126
DB 107 ASNPNTYFPAKRLIGRTWNPDSVOODIKFLPKYVEKTKRYIOVDIGGQTKTFAPE 165
127 ISAWLTFMKKETAAYLGKKTTHAVVTPAYFNDAORATKAGTIAGLNVRITNEPTA 186
DB 166 VSGMLGKMKQIAEDYLGKKTTHAVVTPAYFNDAORATKAGTIAGLNVRITNEPTA 225
QY 197 AAIAYGLDKREGEKNILVFDLGGTFDVSLTIDNGVEEVATNGDTILGGEFDORVME 246
DB 226 AAIAYGLDKREGEKNILVFDLGGTFDVSLTIDNGVEEVATNGDTILGGEFDORVME 285
QY 247 HFIRIKYKKTGKDVAKDRRAVOKLRREVEKAKRAISSOHARIEIESFEGEDPSTILTR 306
DB 286 HFAQIFQKHDVTKNKKAMAKLRREKAKRAISSOHARIEIESFEGEDPSTILTR 345
QY 307 AKFEELNMLDFRSTMKPVOKVLESDLSKSDIDEIVLVGSGTRIPRIQOVLVEFFNGKRP 366
DB 346 AKFEELNMLDFRSTMKPVOKVLESDLSKSDIDEIVLVGSGTRIPRIQOVLVEFFNGKRP 405
QY 367 SSGINDEAVAYGAOVAGVLSGDDTGDVLLDVCPLTGLIETVGVWTKLIPRNTVVP 426
DB 406 SSGINDEAVAYGAOVAGVLSGDDTGDVLLDVCPLTGLIETVGVWTKLIPRNTVVP 465
QY 427 TKKSQIFSTASNOPTVITIKYEGEERPLTKDNHLLGTPOLGIPAPRPRVPOLEVTFEID 486
DB 466 TKKSQIFSTASNOPTVITIKYEGEERPLTKDNHLLGTPOLGIPAPRPRVPOLEVTFEID 525
QY 487 VNGILRVTAEDKGTGNKKNITITNDONRLTPREIERMVNDAKFAEDKKLERIDTRNE 546

DB 526 ANGIILVSATDGTGKSESITIANDKGRISGDDIRMVBEAKKYAEDAKFAKSEARNT 585
QY 547 LESYAVSLKNQIDGKELGKLSSEDEKTEKMEKAEKTEKLESHOD-ADIEDFAKAKKEL 605
DB 586 FENFVHYVKNVSYNG--ELAEIMDEDDKETVLDNVNLSLEMLDNDSDVAEAEDEFEEKMASF 643
QY 606 EEIVOPITISKLYSAG 621
DB 644 KESVEPIIAKASASQ 659

RESULT 7

US-08-797-358B-3
Sequence 3, Application US/08797358B
Patent No. 6268478
GENERAL INFORMATION:
APPLICANT: Adams, John
TITLE OF INVENTION: INTRACELLULAR VITAMIN D BINDING PROTEIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,358B
FILING DATE: 11-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,491
FILING DATE: 12-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-CE 3165
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-797-358B-3

Query Match 59.2%; Score 1938; DB 4; Length 643;
Best Local Similarity 61.7%; Pred. No. 1e-147;
Matches 379; Conservative 107; Mismatches 122; Indels 6; Gaps 5;

QY 13 VGIDIGTTSYSCVGFNGRVEIIANDGNRIIPSYAFTPEGERLIGDAKNOLSNPN 72
DB 9 VGIDIGTTSYSCVGFNGRVEIIANDGNRIIPSYAFT-DTNRLVGDAINQALNPLN 67
QY 73 TVFDKRLIGRTWNPDSVOODIKFLPKYVEKTKRYIOVDIGGQTKTFAPEISAVYL 132
DB 68 TVFDKRLIGRTWNPDSVOODIKFLPKYVEKTKRYIOVDIGGQTKTFAPEISAVYL 126
QY 133 TTKKETAAYLGKKTTHAVVTPAYFNDAORATKAGTIAGLNVRITNEPTAAIAYG 192
DB 127 TTKKETAAYLGKKTTHAVVTPAYFNDAORATKAGTIAGLNVRITNEPTAAIAYG 186

QY 193 LDKR-EGEKNILVFDLGGTDFVSLTTIDNGVFEVAVATNGDTHLGGEDFDQVMEHF1KL 251
DB 187 LDRGAGERNVLIFFDLGGTDFVSLTIDAGVEYKATAGDTHLGGEDFDQVMEHF1KL 246
QY 252 YKKTGKGVKRDNRVQKLRREVEKAKRALSQHOARIEISFEGEDFSELTAKREE 311
DB 247 FRKRHRKDLNWKRLRRLRTACEERAKRTLSSTQATLEIDSLFEGVFPYTSITRAREE 306
QY 312 LNMDFRSMKPVQKLVLEDSDLKSIDIEIVLGGSTRIPKIQOOLVKEFNKEPSRGIN 371
DB 307 LCSDFRSTLEVEKGRDAKIDKAXIHVDVLYGSTRIPKQKLODFNKEKELKSN 366
QY 372 PDEAVAYGAAGVAGVLSGD--ODTGDVLLDVCPLTGLIETVGVWTKLIPRNVTPRK 429
DB 367 PDEAVAYGAAGVAGVAGVLSGD--ODTGDVLLDVCPLTGLIETVGVWTKLIPRNVTPRK 426
QY 430 SOTFTASDNOPTVITKVEGERPLTKDNHLLGTDFDLPAPRAGVPOIETVFEIDVNG 489
DB 427 TOTFTTSDNOGVTIOYEGERAKTKDNHLLGTDFDLPAPRAGVPOIETVFEIDVNG 486
QY 490 ILRVTAEKGTGNKNTTTNDONRLTPEIERMNVDAKFAEDBKLLKERTIDTRNELES 549
DB 487 ILTVATDRSTGKANKITTTNDONRLTPEIERMNVDAKFAEDBKLLKERTIDTRNELES 546
QY 550 YAVSLKNOIGDKREKLGKLSSEDKETMEKAVEKTEWLESHODADIEDFKAKKLEEV 609
DB 547 HVEFHKGLS--QESLRLDKIPKEDRHKVODKCOEVLAMLEHNLADKEEYEHOKRELEQIC 605
QY 610 OPTIKSLVAGAP 623
DB 606 RPIFSRLYGGFCVP 619

RESULT 8
US-08-214-583-2
Sequence 2, Application US/08214583
Patent No. 5627039
GENERAL INFORMATION:
APPLICANT: Petrela-Smith, Olivia
APPLICANT: Madhwa, Renu
TITLE OF INVENTION: MORTALIN AND METHODS FOR DETERMINING
TITLE OF INVENTION: COMPLEMENTATION GROUP ASSIGNMENT OF CANCER CELLS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Ave, NW
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/214,583
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Auerbach, Jeffrey I.
REGISTRATION NUMBER: 32,680
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-383-7451
TELEFAX: 202-383-6610
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

ORIGINAL SOURCE:
ORGANISM: Murine
IMEDIATE SOURCE:
CLONE: mortalin
US-08-214-583-2

Query Match
Best Local Similarity 47.3%; Score 1547.5; DB 1; Length 679;
Matches 333; Conservative 101; Mismatches 163; Indels 53; Gaps 14;

QY 10 GTVVGIDLTGTYSCVGVFNKRYEIIANDQNRIPSYAFTPEGERLLIGDAKNOITSN 69
DB 53 GAVVGIDLTGTYSCVGVFNKRYEIIANDQNRIPSYAFTPEGERLLIGDAKNOITSN 112
QY 70 PENTVADKRLGRMNDNSVOODIKFLPFKYVEKTRKYIYVDIGGQTKTFAPEISA 129
DB 113 PNTPTATRIITGRYDDEVOKTKNVFKIV-RASNDANYEANG--KLYSPQIGA 168
QY 130 NVLTAKKETAEAYLGKRYTHAVTVPAFPNDARQATAGTAGLNVKRIINEPTAAI 189
DB 169 FVLMKKKETAEAYLGKRYTHAVTVPAFPNDARQATAGTAGLNVKRIINEPTAAI 228
QY 190 AYGLDKREGEKNILVFDLGGTDFVSLTTIDNGVFEVAVATNGDTHLGGEDFDQVMEHF1 249
DB 229 AYGLDKRE--DKVIAVYDLDGGTDFDLSILEIYKGVFEVAVATNGDTHLGGEDFDQVMEHF1 287
QY 250 KLYKKRTGKDVKRDNRVQKLRREVEKAKRALSQHOARIEISFEGEDFS-----ET 303
DB 288 KEFRKRTGYDLTKRDNALORVREAEKACELSSVOTDINLP--YLTMDASQPKILNKK 345
QY 304 LTRAKFEELNMDLFRSTMKPVQKLVLEDSDLKSIDIEIVLGGSTRIPKIQOOLVKEFNK 363
DB 346 LTRAKFEELNMDLFRSTMKPVQKLVLEDSDLKSIDIEIVLGGSTRIPKIQOOLVKEFNK 404
QY 364 KEPSRGINDDEAVAYGAAGVAGVLSGDODTGDVLLDVCPLTGLIETVGVWTKLIPRNVTPRK 423
DB 405 RAPSKAVNPDEAVAYGAAGVAGVLSGDODTGDVLLDVCPLTGLIETVGVWTKLIPRNVTPRK 462
QY 424 VPTTKKSQIFSTASDNOPTVITKVEGERPLTKDNHLLGTDFDLPAPRAGVPOIETVFE 483
DB 463 TPTTKKSQIFSTASDNOPTVITKVEGERPLTKDNHLLGTDFDLPAPRAGVPOIETVFE 522
QY 484 ETDVNGILRVTAEKGTGNKNTTTNDONRLTPEIERMNVDAKFAEDBKLLKERTIDTRNELES 541
DB 523 DIDANGIVHVSADKGTGREGQIVL--QSSGGLSKDIDEMVKNVKAEEYAEDEDREKKEVREA 581
QY 542 -----DTRNELESY-----AYSLKNOIGDKREKLGKLSSEDKETMEKAVEEKI 584
DB 582 VNMAEGITHDTETKMEEFKQDLPADCECNKLEISKVRALLAKDSEGTGENIRQAA----- 637
QY 585 EWLSEHODADIEDFKAKKRELEEVQPIISKLYGSAGPPTEG--EDTAE 632
DB 638 -----SLQASAKLFEWAYKKM-----ASERBSG--GSSGTGQKRDONE 676

RESULT 9
US-08-441-139-4
Sequence 4, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltup, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530

QY	251	LYKKETGDKVKNRAVOKLREREYKCKRBAALSSOHORIE---IESYEG-EDPSESLTR	306
Db	240	SFOKETGINTLNDRMAVORVAEAAEKAKIELISTLEITDITLPTFSSOSTGAKHLSLKR	299
QY	307	AKPELNMDEFRSTMKPVOKLVEDSDLK-KSDIDEIYLVGSGTRIPKIOOLYKFEFNGKE	365
Db	300	AKPELVDELIERITLBPCKKALSDAGIKDNSKVDVYLVGSMTRVPRVIGRVDF-FGKE	358
QY	366	PSRGINPDEAVAYCAVAAOVAGVLSDDOTGDVLVLYCPLTGLIETVGVMTKILPRNTVV	425
Db	359	PCOCVNPDEAVAYCAALIOGGILTG--DVDRVLLLDVAPLISGIEITLGVFTPLERNTYI	416
QY	426	PTKKSQITSTASDNQPYTITIVYGEGERPLTRDNHLGLTFDLTGTPRPRGVQIEVFEI	465
Db	417	PTKKSQVSTFEDQOTAVTATITIVYGERKMAIDNMLTLOFSLGPIPRPRGVQIEVFEI	476
QY	486	DVNGILRTATADCKGNGKNKTTINDQNRILPRELIERMVANDAEKFAEDKKLERIDTRN	545
Db	477	DANGIVHVSANDKASGKEQTIKI--OSSGGLSDEIFIKKVMADQADRAEDDEKRKHVELKN	535
QY	546	ELESAYSLKNGIDGKERLKGKLSSEDPKEMKEKVAEKEIML	587
Db	536	SSEGLIHVSKEKLD--YCDKAVGADKSNIESIAIKDLRECL	574

```

11
US-08-472-534-5
: Sequence 5, Application US/08/472534
: Patent No. 5918620
: GENERAL INFORMATION:
:   APPLICANT: Hamel, Josee
:   APPLICANT: Brodeur, Bernard R
:   APPLICANT: Martin, Denis
:   TITLE OF INVENTION: HEAT SHOCK PROTEIN HSP72 FROM
:   TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE
:   NUMBER OF SEQUENCES: 6
:   CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Fish & Neave
:   STREET: 1251 Avenue of the Americas
:   CITY: New York
:   STATE: New York
:   COUNTRY: United States of America
:   ZIP: 10020
:
: * COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/472,534
:   FILING DATE:
:   CLASSIFICATION: 424
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Haley Jr, James F
:     REGISTRATION NUMBER: 27,794
:     REFERENCE/DOCKET NUMBER: Blovac-2
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: 212-596-9000
:       TELEFAX: 212-596-9090
:       TELEXT: 14-8367
:   INFORMATION FOR SEO ID NO: 5:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 607 amino acids
:       TYPE: amino acid
:     TOPOLOGY: linear
:   MOLECULE TYPE: protein
:
: US-08-472-534-5

```

Query Match	44.18;	Score 1444;	DB 2;	Length 607;
Best Local Similarity	49.68;	Pred. No. 5.2e-108;		
Matches 308;	Conservative 99;	Mismatches 170;	Indels 44;	Gaps 11

OY	12	VWIDIGTTCVCCVCENKGRREIITAAINOONIPSPVAFPBEGRLGDAAKNOJNSNE	71
Dd	4	IIGDGLGTNAVAVALBETESKIILANPEGNRTPVSVSf-KNGEI IYGDAKKRAOVANPD	62
OY	72	NTVEDAKRLIGRTNNDSVOODIKFPLPFKVVEKKTKPKPYIOVDIGGGOTKFAPBEISAMV	131
Dd	63	-TVISIRSKMOTS-----EK-----VSANG---KEYTFOEISAMI	93
OY	132	LTKKKEITAELGKKVTHAVVTVAATYRNDOROKTKAGTIAGINVRITINEPTAALAY	151
Dd	94	LOYLKGAVEDLGERKYTRAVTTPAYENDOROKTKAGKIAGLEVRIYNEPTAALAY	153
OY	192	GLDKREGKENLVLDLGGCFEDVSLTLTDNCVFEEVATNTGDHLCGEDPFOORVHEFIKL	251
Dd	154	GLDKTDEKEELVLVDLGGGTFDVSIILELGSDVEDEVLTAGDNKLGGDDFOOKIITHLYAE	213
OY	252	YKKTGKDVKDRKNABOVOKLRREVKEAKRALSSOHARIEIESFEYEB- - -DESETLTRA	307
Dd	214	EKKENGIDISTDKMAQOQLKDAEKAKKDISGVNSTOISLPFTJAGBAPLHLEMLTRA	273
OY	308	KFEELNMDLFSTMKPPQOKVULESDDLKKSIDELVLVGSGSTRIPKIQOLYKEFFNKEPS	367
Dd	274	KFDLTRLVRYTVPVAROALSADGLSISEIDEVILVGGSTRIPAVEAVAKA- ETGKEPN	332
OY	368	RGINDEVAAGAIVONGVLESDDDTGDLVLDCPTLGIETGYGWTLPIPRNVVPT	427
Dd	333	KSNVPDEVVAMGAILOGGVITG--DVMDVLLDVTPLSLGIELMGCVFTLLIBRNTIPT	350
OY	428	KKSQIFASTASDNOPVTVIKYVEGERPLTKDNHLGLTFDLTGIPPARGPVQIEVTEIDV	487
Dd	391	SKSOVFSTAANOQPAVHIHVLOGGERPMAANKTGLRGFOLDIPAAPRGIOIEVTEPIDK	450
OY	488	NGILIRVAEDGCGCNKKKITTTNQNLTPPELIERMVNDAEKPFAEBCKLIKERIDPNEL	547
Dd	451	NGIYSVAKKIDLGTOKEOITVIAQSNG-LTDEIDERMKKDMEANNESDKKKEEZDLNNEY	509
OY	548	ESVAYSLSKNOIGDERKGLOKLSBEDKTEMKAVEEKIEMLESHODAIEDFPKAKKELEE	607
Dd	510	DOAFATEFKETIKEH- - -GKGFDAERDAQAALD- - -DLKAGDNNLDMKAKLEALNE	563
OY	608	IVOPILISKLYSACRPPTGEE	628
Dd	564	KAGGAAVLYEQAAAAQAOAE	564

```

RESULT 12
5240706-1
; Patent No. 5240706
; APPLICANT: FAJLDS, DARYL
; TITLE OF INVENTION: INTRANSSAL ADMINISTRATION OF MYCOPLASMA
; HYDROPHOBIC ANTIGEN
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/334,586
; FILING DATE: 07-APR-1989
; SEQ ID NO.: 1
; LENGTH: 600
5240706-1

```

Query Match	39.8%;	Score 1303.5;	DB 6;	Length 600;
Best Local Similarity	45.38;	Pred. No. 1e-96;		
Matches 282;	Conservative 106;	Mismatches 187;	Indels 47;	Gaps 11

```

0Y 12 VVGIDLTGTYCYVGVKFKGRREITIANDOGNITPSYVAFPFEGEGLRLIGDAKNLQTSNP 71
Db 11 :::::::::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6 ILGIDLGTTNSCAVAIENQKPRVLENPGRKTTTFSYVAF-KNNEIYODAKKROLENP 64
0Y 72 NTVEADAKLIGRTMNDPSVOODIKFLPKVYVEKTKPYIOVDIGGQTKTFPAPEISAMV 133
Db 65 -AIAISIKIKLMQTD-----KTVRANERDVI-----PEISIKI 95

```

TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 52:

Db 121 SEDKEMEKAVEEKIEWLESHQDADIEDFKAKKELEIYQPIISKLYGSACPPPTGEED 180

QY 630 TAE 632
111
DB 181 TAE 183

RESULT 15
US-08-441-139-10
; Sequence 10, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: Wiltrop, Dr. Karl D.
; APPLICANT: Robinson, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,139
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGILLO, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-139-10

Query Match 25.2%; Score 824; DB 1; Length 168;
Best Local Similarity 98.2%; Pred. No. 5.9e-59;
Matches 161; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 469 IPPAPRGVQIEVTFEIDVNGILRTVAEDKGTGNKNTTTNDQNRITPPEIERMVDNAE 528
DB 1 IPPAPRGVQIEVTFEIDVNGILRTVAEDKGTGNKNTTTNDQNRITPPEIERMVDNAE 60
QY 529 KFAEDKKLKERIDTRNELESYAVSLKNOIGDKELGKLSSEDKETMEKAVEEKIEWLE 588
DB 61 KFAEDKKLKERIDTRNELESYAVSLKNOIGDKELGKLSSEDKETMEKAVEEKIEWLE 120
QY 589 SHODADIEDFKAKKKELEIVQPTISKLKSGAGPPTGGEEDTAE 632
DB 121 SHODADIEDFKAKKKELEIVQPTISKLKSGAGPPTGGEEDTAE 164

Search completed: October 9, 2002, 11:17:45
JOB time: 441 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:13:09 ; Search time 34.99 Seconds
(without alignments)
1754.818 Million cell updates/sec

Title: US-09-806-955a-1

Perfect score: 3273

Sequence: 1 MEEDKKEDVGVVIGDGT.....AGPPTGEEDTAELHHHHH 639

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

1 number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR-71:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	98.0	654	2 A27414	dnak-type molecula
2	3201	97.8	655	2 A37048	dnak-type molecula
3	3200	97.8	654	1 HHRTGB	dnak-type molecula
4	3182.5	97.2	653	2 A29821	dnak-type molecula
5	3168	96.8	652	2 I50242	dnak-type molecula
6	2722.5	83.2	667	2 S24782	dnak-type molecula
7	2719.5	83.1	667	2 D44261	dnak-type molecula
8	2710	82.8	656	2 JN0666	dnak-type molecula
9	2618.5	80.0	661	2 T15513	heat shock 70K pro
10	2594.5	79.3	661	2 A32475	dnak-type molecula
11	2559	78.2	657	2 T34037	heat shock 70K pro
12	2356.5	72.0	667	2 S21879	dnak-type molecula
13	2347.5	71.7	668	2 S21880	dnak-type molecula
14	2332.5	71.3	668	2 T04080	dnak-type molecula
15	2330	71.3	668	2 T46574	dnak-type molecula
16	2322.5	71.0	663	2 T04078	dnak-type molecula
17	2305.5	70.4	663	2 T03581	dnak-type molecula
18	2298	70.2	668	2 S71171	dnak-type molecula
19	2285	69.8	672	2 T43723	dnak-type molecula
20	2278	69.6	672	2 T43716	dnak-type molecula
21	2261	69.1	642	2 T45298	dnak-type molecula
22	2240.5	68.5	655	2 T06598	dnak-type molecula
23	2237	68.3	655	2 H86222	hypothetical prote
24	2231	68.2	658	2 S38890	dnak-type molecula
25	2227.5	68.1	663	2 T50464	glucose-regulated
26	2216	67.7	663	2 T38155	78 kD glucose regu
27	2214.5	67.4	608	2 T05741	dnak-type molecula
28	2205	67.4	666	2 T06358	dnak-type molecula
29	2174	66.4	682	1 HHBYK2	dnak-type molecula

30	2156	65.9	663	2 S20877	dnak-type molecula
31	2108	64.4	651	2 JC7132	heat shock protein
32	2100	64.2	655	2 A48468	dnak-type molecula
33	2083	63.6	646	2 S07197	dnak-type molecula
34	2083	63.6	646	2 JC4853	dnak-type molecula
35	2083	63.6	646	2 A27077	dnak-type molecula
36	2080	63.6	641	2 PC7036	heat shock protein
37	2079	63.5	647	2 T41121	heat shock protein
38	2079	63.5	650	2 S11456	dnak-type molecula
39	2078	63.5	645	2 S41372	dnak-type molecula
40	2077	63.5	646	2 A45935	dnak-type molecula
41	2076.5	63.4	640	2 S37394	dnak-type molecula
42	2074.5	63.4	648	2 S53498	dnak-type molecula
43	2074.5	63.4	656	2 S51712	dnak-type molecula
44	2074	63.4	646	2 A35922	dnak-type molecula
45	2072	63.3	639	2 S20139	dnak-type molecula

ALIGNMENTS

RESULT 1

A27414
dnak-type molecular chaperone GRP78 precursor - Chinese hamster
N:Alternate names: glucose-regulated 78K protein; Ig heavy chain-binding protein
C:Species: Cricetus griseus (Chinese hamster)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999
C:Accession: A27414
R:Ring, J.; Wooden, S.K.; Ritz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.
Gene 55, 147-152, 1987
A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protei
A:Reference number: A27414; MUID:87305586
A:Accession: A27414
A:Molecule type: mRNA
A:Residues: 1-654 <TIN>
A:Cross-references: GB:M17169; NID:g191090; PIDN:AAA51448.1; PID:g304510
C:Genetics:
A:Gene: GRP78
C:Function:
A:Description: Involved in protein folding and assembling/dissassembling of protein co
C:Superfamily: heat shock protein 70
C:Keywords: ATP; heterotetramer; immunoglobulin; molecular chaperone
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-654/Product: Ig heavy chain binding protein #status predicted <MAT>
F:651-654/Region: endoplasmic reticulum retention signal

Query Match 98.0%; Score 3209; DB 2; Length 654;
Best Local Similarity 99.7%; Pred. No. 8.1e-155;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	2	EEDKKEDVGVVIGDGTTSVCVGFKNRGVEITANDOGNRTPSYVAFPEGRRLGDA	61
DB	20	EEDKKEDVGVVIGDGTTSVCVGFKNRGVEITANDOGNRTPSYVAFPEGRRLGDA	79
QY	62	AKNOLTSNPENTVFDARLIGRTWMDPSVODIKFLPKVVEKTKPIVDIGGGGQTK	121
DB	80	AKNOLTSNPENTVFDARLIGRTWMDPSVODIKFLPKVVEKTKPIVDIGGGGQTK	139
QY	122	FAPEISAMVLTAKKETAEAYLGKVTTHAVTVYAFNDARQATKAGTAGLVNRII	181
DB	140	FAPEISAMVLTAKKETAEAYLGKVTTHAVTVYAFNDARQATKAGTAGLVNRII	199
QY	182	NEPFAAAIANGLDKREGEKNITLVDDGGTFDVSLITDNGVEFVATNGDTHLGDEDF	241
DB	200	NEPFAAAIANGLDKREGEKNITLVDDGGTFDVSLITDNGVEFVATNGDTHLGDEDF	259
QY	242	QRVNEHFIKLYKKTKGDKVRKDNRAVOKLRREVKKARALSSQARIEISFEEDGFS	301
DB	260	QRVNEHFIKLYKKTKGDKVRKDNRAVOKLRREVKKARALSSQARIEISFEEDGFS	319
QY	302	ETLRAKFEELNMDLFRSTMKPVOKLEDSDLKSDIDEIVLVGSTRIPKIQOLVREFF	361

Db 320 ETLFRANFEEINMDLFRSTMKPVQKVLIEDSLKSDIDEIVLVCSTRIPIQOLVKEEF 379
Qy 362 NGKEPSRCINPDEAVAAVGAAGVLSGDQDPTGLVDVCPPLTGIETVGVMTKLIPR 421
Db 380 NGKEPSRCINPDEAVAAVGAAGVLSGDQDPTGLVDVCPPLTGIETVGVMTKLIPR 439
Qy 422 NTVPPTKKSQIFSTASDNPVTITIKYGEGERPLTKDNHLCTFDLTGTPAPRGVPOIEV 481
Db 440 NTVPPTKKSQIFSTASDNPVTITIKYGEGERPLTKDNHLCTFDLTGTPAPRGVPOIEV 499
Qy 482 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIERVAVDAEFAEDKKLKERI 541
Db 500 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIERVAVDAEFAEDKKLKERI 559
Qy 542 DTRNELESYASLKNQIDGKELGKLSSEDEKTEKAVEEKIEMWLESODADIEDFKAK 601
Db 560 DTRNELESYASLKNQIDGKELGKLSSEDEKTEKAVEEKIEMWLESODADIEDFKAK 619
Qy 602 KKELEIYQPIISKLKYSAGPPTGEEDTAE 632
620 KKELEIYQPIISKLKYSAGPPTGEEDTSE 650

RESULT 2
A37048
dnak-type molecular chaperone grp78 precursor - mouse
N:Alternate names: B1P; glucose-regulated 78k protein; Ig heavy chain-binding protein
C:Species: Mus musculus (house mouse)
C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 20-Aug-1999
C:Accession: A37048; B37048; J00094; I49137; A31934
R:Kozutsumi, Y.; Normington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.
J. Cell Sci. Suppl. 11, 115-137, 1989
A:Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulat
A:Reference number: A37048; MUID:90130686
A:Accession: A37048
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-655 <K21>
A:Accession: B37048
A:Molecule type: protein
A:Residues: 20-41, 'X', 43-46 <K22>
R:Parfett, C.L.J.; Hofbauer, R.; Brudzynski, K.; Edwards, D.R.; Denhardt, D.T.
Gene 82, 291-303, 1989
A:Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero
A:Reference number: J00094; MUID:90060818
A:Accession: J00094
A:Molecule type: mRNA
A:Residues: 488-590 'G', 592-595, 'E', 597-655 <PAR>
R:References: GB:M30779; NID:g193644; PIDN:AAA37742.1; PID:g193645
Gene 158, 225-229, 1995
R:Allman, J.B.; Mote, P.L.; Walford, R.L.; Spindler, S.R.
A:Title: Structure and regulation of the mouse GRP78 (B1P) promoter by glucose and calci
A:Reference number: I49137; MUID:95331621
A:Accession: I49137
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28 <RES>
A:Cross-references: EMBL:U016277; NID:g829364; PIDN:AAA6734.1; PID:g829365
R:Haas, I.G.; Moe, T.
Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988
A:Title: cDNA cloning of the immunoglobulin heavy chain binding protein.
A:Reference number: A31934; MUID:8816922
A:Accession: A31934
A:Molecule type: mRNA
A:Residues: 514-589 'D', 591-595, 'E', 597-655 <HAA>
A:Cross-references: GB:M19351; NID:g192231; PIDN:AAA37315.1; PID:g387113
C:Genetics: grp78
A:Gene: grp78
C:Function:
A:Description: Involved in protein folding and assembling/disassembling of protein comp
C:Superfamily: heat shock protein 70
C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon
F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-655/Product: dnak-type molecular chaperone grp78 #status experimental <MAT>
F:652-655/Region: endoplasmic reticulum retention signal

Query Match 97.8%; Score 3201; DB 2; Length 655;
Best Local Similarity 99.4%; Pred. No. 2,1e-134;
Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EEDKKEDGVTVGIDLGITVSCVGFKNRVEIINDDGNRITPSVAFTEGGERLIGDA 61
Db 21 EEDKKEDGVTVGIDLGITVSCVGFKNRVEIINDDGNRITPSVAFTEGGERLIGDA 80
Qy 62 AKNOLTSNPENTVFDKRLIGRTMNDPSVQODIKFLPKRVVEKTKPYIQVDIGGQRT 121
Db 81 AKNOLTSNPENTVFDKRLIGRTMNDPSVQODIKFLPKRVVEKTKPYIQVDIGGQRT 140
Qy 122 FAPETISAMVLTMKKETAEATIGKRVTAHVTVVPAFMDAQRAKKDAGTLAGLVNMI 181
Db 141 FAPETISAMVLTMKKETAEATIGKRVTAHVTVVPAFMDAQRAKKDAGTLAGLVNMI 200
Qy 182 NEPTAAATAYGLDKREGKNTLVFDLGGTDPVSLTIDNGVEFVATNGDTHLGEDPD 241
Db 201 NEPTAAATAYGLDKREGKNTLVFDLGGTDPVSLTIDNGVEFVATNGDTHLGEDPD 260
Qy 242 QRVMEFTIKLYKKKTGDKVRKDNRAVOKLRREVEKAKRALSQHQARIESEFVEGDFS 301
Db 261 QRVMEFTIKLYKKKTGDKVRKDNRAVOKLRREVEKAKRALSQHQARIESEFVEGDFS 320
Qy 302 ETLFRANFEEINMDLFRSTMKPVQKVLIEDSLKSDIDEIVLVCSTRIPIQOLVKEEF 361
Db 321 ETLFRANFEEINMDLFRSTMKPVQKVLIEDSLKSDIDEIVLVCSTRIPIQOLVKEEF 380
Qy 362 NGKEPSRCINPDEAVAAVGAAGVLSGDQDPTGLVDVCPPLTGIETVGVMTKLIPR 421
Db 381 NGKEPSRCINPDEAVAAVGAAGVLSGDQDPTGLVDVCPPLTGIETVGVMTKLIPR 440
Qy 422 NTVPPTKKSQIFSTASDNPVTITIKYGEGERPLTKDNHLCTFDLTGTPAPRGVPOIEV 481
Db 441 NTVPPTKKSQIFSTASDNPVTITIKYGEGERPLTKDNHLCTFDLTGTPAPRGVPOIEV 500
Qy 482 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIERVAVDAEFAEDKKLKERI 541
Db 501 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIERVAVDAEFAEDKKLKERI 560
Qy 542 DTRNELESYASLKNQIDGKELGKLSSEDEKTEKAVEEKIEMWLESODADIEDFKAK 601
Db 561 DTRNELESYASLKNQIDGKELGKLSSEDEKTEKAVEEKIEMWLESODADIEDFKAK 620
Qy 602 KKELEIYQPIISKLKYSAGPPTGEEDTAE 632
Db 621 KKELEIYQPIISKLKYSAGPPTGEEDTSE 651

RESULT 3
HHR7CB
dnak-type molecular chaperone precursor - rat
N:Alternate names: B1P; glucose-regulated 78k protein; Ig heavy chain-binding protein
N:Contains: steroidogenesis-activator polypeptide
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-Jun-1999
C:Accession: A23948; A60134; A26257
R:Munro, S.; Pelham, H.R.B.
Cell 46, 291-300, 1986
A:Title: An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated p
A:Reference number: A23948; MUID:86245075
A:Accession: A23948
A:Molecule type: mRNA
A:Residues: 1-654 <MUN>
A:Cross-references: GB:M14050; NID:g203150; PIDN:AAA40817.1; PID:g203151
R:Pedersen, R.C.; Browne, A.C.
Science 236, 188-190, 1987
A:Title: Steroidogenesis-activator polypeptide isolated from a rat leydig cell tumor.
A:Reference number: A60134; MUID:87177981

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:17:49 ; Search time 19.71 seconds
(without alignments)
1255.291 Million cell updates/sec

Title: US-09-806-955A-1
Perfect score: 3273
Sequence: 1 MEEDKEDVGTVGIDLCTT.....AGPPTGEDETAELHHHHH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

1 number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt-40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3216	98.3	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	98.0	654	1 GR78_MESAU	P07823 mesocricetu
3	3205	97.9	655	1 GR78_MOUSE	P20029 mus musculu
4	3200	97.8	654	1 GR78_RAT	P06761 rattus norv
5	3168	96.8	652	1 GR78_CHICK	Q90593 gallus gall
6	3069	93.8	658	1 GR78_XENLA	Q91883 xenopus lae
7	2722.5	83.2	667	1 GR78_APLCA	Q16956 aplysia cal
8	2708	82.7	656	1 HS7C_DROME	P29844 drosophila
9	2594.5	79.3	661	1 HS7C_CAEEL	P27420 caenorhabdi
10	2356.5	72.0	667	1 BIP4_TOBAC	Q03684 nicotiana t
11	2347.5	71.7	668	1 BIP5_TOBAC	Q03685 nicotiana t
12	2332.5	71.3	663	1 BIP3_MAIZE	Q24581 zea mays (m
13	2328.5	71.1	666	1 BIP_LYCES	P49118 lycopersico
14	2328.5	71.1	668	1 BIP_SPIOL	Q42434 spinacia ol
15	2322.5	71.0	663	1 BIP2_MAIZE	P24067 zea mays (m
16	2301	70.3	669	1 BIP1_ARATH	Q91K73 arabidopsis
17	2299	70.2	668	1 BIP2_ARATH	Q39043 arabidopsis
18	2227.5	68.1	662	1 GR78_NEUCR	P78695 neurospora
19	2216	67.7	663	1 GR78_SCHPO	P36604 schizosacch
20	2195	67.1	670	1 GR78_YARLI	Q09170 yarrowia li
21	2174	66.4	682	1 GR78_YEAST	P16474 saccharomyc
22	2100	64.2	655	1 GR78_PLAFO	Q05866 plasmodium
23	2085.5	63.7	641	1 HS71_YEAST	P10591 saccharomyc
24	2083	63.6	646	1 HS7C_HUMAN	P11142 homo sapien
25	2083	63.6	646	1 HS7C_MOUSE	P08109 mus musculu
26	2079	63.5	650	1 HS7C_BOVIN	P08109 mus musculu
27	2078	63.5	644	1 HS71_PICAN	P53421 picchia angu
28	2076.5	63.4	640	1 HS7C_DICDI	P36415 dictyostell
29	2074.5	63.4	655	1 HS71_CANAL	P41397 candida alb
30	2074	63.4	646	1 HS7C_CRICK	P19378 cricetus
31	2072	63.3	638	1 HS72_YEAST	P10592 saccharomyc
32	2065.5	63.1	679	1 GR78_KULFA	P22010 kluyveromyc
33	2065	63.1	648	1 HS71_PUCGR	Q01877 puccinia gr

34	2064	63.1	641	1 HS74_YEAST	P22202 saccharomyc
35	2063.5	63.0	649	1 HS73_YEAST	P09435 saccharomyc
36	2060.5	63.0	646	1 HS70_NEUCR	Q01223 neurospora
37	2060	62.9	643	1 HS71_SCHPO	Q10285 schizosacch
38	2059	62.9	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	62.7	651	1 HS7C_PETRY	P09189 petunia hyb
40	2052	62.7	652	1 HS7D_MANSE	Q9u639 manduca sex
41	2049	62.6	651	1 HS70_ONCMY	P08108 oncorhynch
42	2046	62.5	641	1 HS72_BOVIN	Q27965 bos taurus
43	2045	62.5	651	1 HS71_ARATH	P22953 arabidopsis
44	2043.5	62.4	639	1 HS72_HUMAN	P54652 homo sapien
45	2039.5	62.3	641	1 HS71_HUMAN	P08107 homo sapien

ALIGNMENTS

RESULT	ID	STANDARD	PRT	554 AA.
GR78_HUMAN				
AC	P11021; Q9NPF1;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin heavy chain binding protein) (BIP) (Endoplasmic reticulum lumenal Ca2+ binding protein grp78).			
DE	binding protein grp78).			
GN	HSP45 OR GRP78.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88283347; PubMed=2840249;			
RA	ting J., Lee A.S.;			
RT	"Human gene encoding the 78,000-dalton glucose-regulated protein and its pseudogene: structure, conservation, and regulation.";			
RL	DNA 7:275-286(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cervical carcinoma;			
RA	Chao C.C.K.;			
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fibroblast;			
RA	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;			
RT	"GRP78 is involved in the quality control of the LDL-receptor.";			
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Bermudez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;			
RT	"Sequence differences between human grp78/BIP isolated from HeLa cells and previously reported human sequences.";			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 1-25 FROM N.A.			
RX	MEDLINE=93126087; PubMed=1480470;			
RA	Chao C.C.K., Lin-Chao S.;			
RT	"A direct-repeat sequence of the human BIP gene is required for A23187-mediated inducibility and an inducible nuclear factor binding.";			
RL	Nucleic Acids Res. 20:6481-6485(1992).			
RN	[6]			
RP	SEQUENCE OF 22-38.			
RC	TISSUE=Breast carcinoma;			
RX	MEDLINE=97295304; PubMed=9150946;			
RA	Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.,			
RT	Stimpson R.J., Dorow D.S.;			
RT	"Two-dimensional electrophoretic analysis of human breast carcinoma proteins: mapping of proteins that bind to the SH3 domain of mixed lineage kinase MLK2.";			

```

RL Electrophoresis 18:588-598(1997).
RN
RC TISSUE-COLON CARCINOMA:
RX MEDLINE-97295306; PubMed-9150948;
RA JI H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
CC
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: M19645; AAA52614.1; -.
DR EMBL; X87949; CAA61201.1; -.
DR EMBL; AJ271729; CAB71335.1; -.
DR EMBL; AF216292; AAF42836.1; -.
DR EMBL; X59969; CAA42595.1; -.
DR PIR; A29821; A29821.
DR HSSP; P19120; 3HSC.
DR SWISS-2DPAGE; P11021; HUMAN.
DR PMMA-2DPAGE; P11021; -.
DR PICT-2DPAGE; P11021; -.
DR Sien-2DPAGE; P11021; -.
DR MIM; 138120; -.
DR InterPro; IPR000866; ER_target.
DR InterPro; IPR001023; HSP70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
KW APP-binding; Endoplasmic reticulum; Signal.
FT SIGNAL 1 18
FT CHAIN 19 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT SITE 651 654 PREVENT SECRETION FROM ER.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
FT SEQUENCE 654 AA; 72333 MW; 59B7DBD85BC32A00 CRC64;
Query Match 98.3%; Score 3216; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 3, 2e-152;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 260 QVMEHFILKLYKKTKGVKRDNRVAVOKLRREVEKAKRALSSOARIESFYEGEDFS 319
QY 302 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIYLVGSTRIPRIQOLVKEEF 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIYLVGSTRIPRIQOLVKEEF 379
QY 362 NGKERSRCINPDEAVAAVQAGVLSGDODPTGLVLLDVCPLTGLITVGVMTKLPR 421
DB 380 NGKERSRCINPDEAVAAVQAGVLSGDODPTGLVLLDVCPLTGLITVGVMTKLPR 439
QY 422 NTVPPTKSOIFSTASDNOPTVTIKVYGERPLTKDNHLGTFDLTGIPAPRGVPOIEV 481
DB 440 NTVPPTKSOIFSTASDNOPTVTIKVYGERPLTKDNHLGTFDLTGIPAPRGVPOIEV 499
QY 482 TFEIDVNGILVTADKDKTKGNKNTITNDONRLPEEIERVNDAEKFAEDKKLKERI 541
DB 500 TFEIDVNGILVTADKDKTKGNKNTITNDONRLPEEIERVNDAEKFAEDKKLKERI 559
QY 542 DTRNELESYAVSLKNQIGDKLSEDEKETEWEKAVEEKIEWLSEHODADIEDFRK 601
DB 560 DTRNELESYAVSLKNQIGDKLSEDEKETEWEKAVEEKIEWLSEHODADIEDFRK 619
QY 602 KKELEIVQPIISKLKYSAGAPPTGEEDTAE 632
DB 620 KKELEIVQPIISKLKYSAGAPPTGEEDTAE 650
RESULT 2
GR78.MESAU STANDARD: PRT: 654 AA.
ID GR78.MESAU
AC P07823;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin
DE heavy chain binding protein) (BIP).
GN HSP45 OR GRP78.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
CX NCBI_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87305586; PubMed-3623104;
RA Ting J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
RT "The nucleotide sequence encoding the hamster 78-kDa
RT glucose-regulated protein (GRP78) and its conservation between
RT hamster and rat.";
RL Gene 55:147-152(1987).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: M17169; AAA51448.1; -.
DR PIR; A27414; A27414.
DR HSSP; P19120; 3HSC.
DR InterPro; IPR000886; ER_target.
DR InterPro; IPR001023; HSP70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00297; HSP70_1; 1.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:15:59 ; Search time 50.75 Seconds

(Without alignments)
2178.202 Million cell updates/sec

Title: US-09-806-955a-1

Sequence: 1 MEEDKKEVGVVIGIDICTT.....AGPPTGSDTAELHHHHH 639

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues 562222

1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	4	09UK02 homo sapien
2	3199	97.7	655	11	09DC41 mus musculu
3	2994	91.5	658	13	091688 xenopus lae
4	2725	83.3	658	5	076180 bombyx mori
5	2717	83.0	656	5	09YVU3 drosophila
6	2607.5	79.7	656	5	062568 suberites d
7	2559	78.2	657	5	002085 caenorhabdi
8	2534	77.4	651	5	024798 echinococu
9	2525.5	77.2	649	5	024895 echinococu
10	2460	75.2	648	5	045038 schistosoma
11	2345.5	71.7	665	10	09M4E8 cucumis sat
12	2330	71.2	668	10	022639 glycine max
13	2325.5	71.1	557	5	0961X6 trichinella
14	2313.5	70.7	675	10	040924 pseudotsuga
15	2305.5	70.4	663	10	024182 oryza sativ
16	2292	70.0	668	10	09FSY7 corylus ave

17	2288.5	69.9	665	10	094IK4	094IK4 scherffelia
18	2285	69.8	672	3	014453	014453 aspergillus
19	2278	69.6	672	3	013280	013280 aspergillus
20	2273	69.4	668	5	09U540	09U540 toxoplasma
21	2263	69.1	669	3	09C1C1	09C1C1 cronartium
22	2261	69.1	642	5	061001	061001 toxoplasma
23	2246.5	68.6	672	3	09UWE3	09UWE3 aspergillus
24	2243	68.5	659	10	041074	041074 phaeodactyl
25	2240.5	68.5	664	10	039830	039830 glycine max
26	2237	68.3	655	10	004022	004022 arabidopsis
27	2231	68.2	658	10	012752	012752 phytophthor
28	2229.5	68.1	701	5	024928	024928 elmeria ten
29	2214.5	67.7	608	10	040058	040058 hordeum vul
30	2205	67.4	656	3	039804	039804 glycine max
31	2188.5	66.9	665	3	P87028	P87028 pneumocysti
32	2160.5	66.0	665	3	09HG01	09HG01 pichia angu
33	2149.5	65.7	669	10	094H00	094H00 oryza sativ
34	2139.5	65.4	652	3	000053	000053 pneumocysti
35	2116.5	64.7	652	5	007615	007615 plasmodium
36	2108	64.4	651	3	09UW1	09UW1 rhizopus st
37	2093	63.9	659	5	09XZJ2	09XZJ2 crassostrea
38	2080	63.6	642	3	09UW0	09UW0 rhizopus st
39	2079	63.5	647	3	059855	059855 schizosach
40	2076	63.4	661	5	077164	077164 entamoeba h
41	2074.5	63.4	648	10	040980	040980 pisum sativ
42	2069	63.2	646	13	073885	073885 gallus gall
43	2065.5	63.1	647	3	094104	094104 pneumocysti
44	2064.5	63.1	645	3	094106	094106 pneumocysti
45	2057	62.8	652	5	026924	026924 trypanosoma

ALIGNMENTS

RESULT	ID	09UK02	PRELIMINARY:	PRT:	639 AA.
AC	09UK02	01-MAY-2000 (TREMURel. 13, Created)			
DT	01-MAY-2000 (TREMURel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMURel. 19, Last annotation update)				
DE	BIP PROTEIN (FRAGMENT).				
GN	HSP45.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=ARTICULAR CARTILAGE;				
RA	File M.S.;				
RT	"BIP: an Autoantigen associated with Rheumatoid Arthritis.";				
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.				
CC	-1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.				
DR	EMBL: AF188611; AAF13605.1; "				
DR	HSP: P19120; 3HSC.				
DR	InterPro: IPR001023; HSP70.				
DR	Pfam: PF00012; HSP70; 1.				
DR	PRINTS: PR00301; HEATSHOCK70.				
DR	PROSITE: PS00297; HSP70_1; 1.				
DR	PROSITE: PS00329; HSP70_2; 1.				
DR	PROSITE: PS01036; HSP70_3; 1.				
KW	ATP-binding.				
FT	NON_TER	639			
SQ	SEQUENCE	639 AA;	70931 MW;	3421208FF128FD5D CRC64;	
Query Match	100.0%;	Score 3273;	DB 4;	Length 639;	
Best Local Similarity	100.0%;	Pred. No. 3,6e-157;			
Matches 639;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
1 MEEDKKEVGVVIGIDICTTSCGVFNGKRGVEITANDGRRITPSVYAFPEGEIRLGD 60					

Db 1 MEEDKKEDVGVIGIDLGTTYSVGVFKNGRVEIILANDQGNRITPSYVAFTEPGEERLIGD 60
 QY 61 AAKNQLTSPNPTVFAKRLIGRTMNDPSVOODIKFLPFKYVEKTKRPYIOVDIGGQTK 120
 Db 61 AAKNQLTSPNPTVFAKRLIGRTMNDPSVOODIKFLPFKYVEKTKRPYIOVDIGGQTK 120
 QY 121 TFAPEEISAMVLTAKMETAEAYLGGKVTTHAVYVPAYFNDQROATKDGATAGLVNRI 180
 Db 121 TFAPEEISAMVLTAKMETAEAYLGGKVTTHAVYVPAYFNDQROATKDGATAGLVNRI 180
 QY 181 INEPTAAIAVGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFVYVATNDTHLGGEDF 240
 Db 181 INEPTAAIAVGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFVYVATNDTHLGGEDF 240
 QY 241 DQRMENHFKLYKKKTKGDKVKKDNRAVOKLRREVEKAKRALSSQOARIEIESFEEDGF 300
 Db 241 DQRMENHFKLYKKKTKGDKVKKDNRAVOKLRREVEKAKRALSSQOARIEIESFEEDGF 300
 QY 301 SETLTRAKEELNMDLFRSTMKPVOKVLESDDLKKSIDELIVLGGSTRIPKIQOLVKEF 360
 Db 301 SETLTRAKEELNMDLFRSTMKPVOKVLESDDLKKSIDELIVLGGSTRIPKIQOLVKEF 360
 QY 361 FNGKEPSRGIPDEAVAGAAGVAGVLSGDDQDGLVLDVCPRLTIGIETVGVMTKLIP 420
 Db 361 FNGKEPSRGIPDEAVAGAAGVAGVLSGDDQDGLVLDVCPRLTIGIETVGVMTKLIP 420
 QY 421 RNTVPTKKSQIFSTASDNQPTVTIKYEGEERPLTKDNHLGTFDLTGIPPARGVPOIE 480
 Db 421 RNTVPTKKSQIFSTASDNQPTVTIKYEGEERPLTKDNHLGTFDLTGIPPARGVPOIE 480
 QY 481 VTFFIDVNGILRVTAEDKGTGNKKNITITNDQNRILPPEIERMVNDAEKFAEDKKLER 540
 Db 481 VTFFIDVNGILRVTAEDKGTGNKKNITITNDQNRILPPEIERMVNDAEKFAEDKKLER 540
 QY 541 IDTRNELESYVSLKNOGDKKLGKLSSEDKETMEKAVEEKIEMWESHODADIEDKKA 600
 Db 541 IDTRNELESYVSLKNOGDKKLGKLSSEDKETMEKAVEEKIEMWESHODADIEDKKA 600
 QY 601 KKELEEIYOPILSKLYGSAGRPPTGEEDTAELNHNHNH 639
 Db 601 KKELEEIYOPILSKLYGSAGRPPTGEEDTAELNHNHNH 639
 RESULT 2
 Q9DC41 PRELIMINARY: PRT: 655 AA.
 AC 09DC41
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 AC 09DC41: 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HEAT SHOCK 70KD PROTEIN 5 (GLUCOSE-REGULATED PROTEIN, 78KD).
 GN HSPA5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=1090;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pease G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Buit C., Fletcher C., Fujita K., Gariboldi M.,
 RA Guelinckich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilmink L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 DR EMBL: AK004578; BAB23387.1; -.
 DR HSSP: P19120; 3HSC.
 DR MGD: MGI:95835; Hspa5.
 DR InterPro: IPR000886; ER_target.
 DR InterPro: IPR001023; HSP70.
 DR Pfam: PF00012; HSP70; 1.
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
 KW ATP-binding.
 SQ SEQUENCE 655 AA: 72422 MW: 69620CA2C896DE6C CRC64:
 Query Match 97.7%; Score 3199; DB 11; Length 655;
 Best Local Similarity 99.2%; Pred. No. 2e-153;
 Matches 626; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EEDKKEDVGVIGIDLGTTYSVGVFKNGRVEIILANDQGNRITPSYVAFTEPGEERLIGD 61
 Db 21 EEDKKEDVGVIGIDLGTTYSVGVFKNGRVEIILANDQGNRITPSYVAFTEPGEERLIGD 80
 QY 62 AAKNQLTSPNPTVFAKRLIGRTMNDPSVOODIKFLPFKYVEKTKRPYIOVDIGGQTK 121
 Db 81 AAKNQLTSPNPTVFAKRLIGRTMNDPSVOODIKFLPFKYVEKTKRPYIOVDIGGQTK 140
 QY 122 FAPBEISAMVLTAKMETAEAYLGGKVTTHAVYVPAYFNDQROATKDGATAGLVNRI 181
 Db 141 FAPBEISAMVLTAKMETAEAYLGGKVTTHAVYVPAYFNDQROATKDGATAGLVNRI 200
 QY 182 NEPTAAIAVGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFVYVATNDTHLGGEDF 241
 Db 201 NEPTAAIAVGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFVYVATNDTHLGGEDF 260
 QY 242 QRMVHFILKLYKKKTKGDKVKKDNRAVOKLRREVEKAKRALSSQOARIEIESFEEDGF 301
 Db 261 QRMVHFILKLYKKKTKGDKVKKDNRAVOKLRREVEKAKRALSSQOARIEIESFEEDGF 320
 QY 302 ETLTRAKEELNMDLFRSTMKPVOKVLESDDLKKSIDELIVLGGSTRIPKIQOLVKEF 361
 Db 321 ETLTRAKEELNMDLFRSTMKPVOKVLESDDLKKSIDELIVLGGSTRIPKIQOLVKEF 380
 QY 362 NGKEPSRGIPDEAVAGAAGVAGVLSGDDQDGLVLDVCPRLTIGIETVGVMTKLIP 421
 Db 381 NGKEPSRGIPDEAVAGAAGVAGVLSGDDQDGLVLDVCPRLTIGIETVGVMTKLIP 440
 QY 422 NTVPPTKKSQIFSTASDNQPTVTIKYEGEERPLTKDNHLGTFDLTGIPPARGVPOIEV 481
 Db 441 NTVPPTKKSQIFSTASDNQPTVTIKYEGEERPLTKDNHLGTFDLTGIPPARGVPOIEV 500
 QY 482 TFEIDVNGILRVTAEDKGTGNKKNITITNDQNRILPPEIERMVNDAEKFAEDKKLERI 541
 Db 501 TFEIDVNGILRVTAEDKGTGNKKNITITNDQNRILPPEIERMVNDAEKFAEDKKLERI 560
 QY 542 DTRNELESYVSLKNOGDKKLGKLSSEDKETMEKAVEEKIEMWESHODADIEDFAK 601
 Db 561 DTRNELESYVSLKNOGDKKLGKLSSEDKETMEKAVEEKIEMWESHODADIEDFAK 620
 QY 602 KKELEEIYOPILSKLYGSAGRPPTGEEDTAELNHNHNH 632
 Db 621 KKELEEIYOPILSKLYGSAGRPPTGEEDTAELNHNHNH 651
 RESULT 3
 Q91688 PRELIMINARY: PRT: 658 AA.
 AC Q91688

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:04:54 ; Search time 58.8 Seconds
(without alignments)
1207.077 Million cell updates/sec

Title: US-09-806-955a-1
Perfect score: 3273
Sequence: 1 MEEDKEDVGTWVGIDLGTT.....AGPPPTGEDTALHHHHH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
11 number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	21	AAV90693 639 residue human
2	3225	98.5	633	21	AAV90694 633 residue human
3	3209	98.0	654	15	AAE53075 Hamster GRP78. Cr
4	3182.5	97.2	653	22	AAE12985 Human Hsp70 family
5	3182.5	97.2	653	22	AAE12985 Human heat shock p
6	3171	96.9	656	14	AAE24934 Human glucose regu
7	3168	96.8	656	22	AAE53076 Chicken GRP78. Ga
8	2717	83.0	656	22	AAE58666 Drosophila melanog
9	2299	70.2	668	21	AAE41382 Arabidopsis thalia
10	2299	70.2	691	21	AAE41381 Arabidopsis thalia
11	2258.5	69.0	642	21	AAE39285 Arabidopsis thalia

12	2258.5	69.0	658	21	AAE39284 Arabidopsis thalia
13	2237	68.3	655	21	AAE30912 Arabidopsis thalia
14	2174	66.4	682	22	AAE70749 S cerevisiae apopt
15	2125.5	64.9	687	22	AAE70875 C albicans apoptos
16	2083	63.6	646	19	AAE54364 Heat shock cognate
17	2083	63.6	646	20	AAE17407 Human heat shock c
18	2083	63.6	646	20	AAE17408 Mouse heat shock c
19	2083	63.6	646	21	AAE23649 Mouse heat shock p
20	2083	63.6	646	21	AAE23651 Human heat shock p
21	2083	63.6	646	22	AAE12987 Human Hsp70 family
22	2083	63.6	646	22	AAE82535 Human heat shock p
23	2083	63.6	890	21	AAE22938 GFP-HSC70 fusion p
24	2074.5	63.4	656	18	AAE01638 Candida albicans h
25	2064	63.1	645	11	AAE03927 Rat HSP (rethsp70)
26	2041	62.4	647	11	AAE03928 Xenopus laevis HSP
27	2039.5	62.3	624	21	AAE23252 Human Hsp72 (heat
28	2039.5	62.3	641	19	AAE54349 Human heat shock 7
29	2039.5	62.3	641	21	AAE23652 Human heat shock p
30	2039.5	62.3	641	22	AAE12986 Human Hsp70 family
31	2039.5	62.3	641	22	AAE82534 Human heat shock p
32	2036.5	62.2	646	21	AAE53604 Arabidopsis thalia
33	2035.5	62.1	633	14	AAE43002 Mouse SLPI homolo
34	2032.5	62.1	647	20	AAE31380 T. gondii antigen
35	2032.5	62.1	647	22	AAE49099 Toxoplasma gondii
36	2032	62.1	634	11	AAE03930 Gallus gallus HSP
37	2030	62.0	651	22	AAE60514 Drosophila melanog
38	2021	61.7	640	21	AAE23653 Human heat shock p
39	2017.5	61.6	640	11	AAE03929 Homo sapiens HSP (
40	2014	61.5	642	21	AAE23650 Rat heat shock pro
41	2013	61.5	640	18	AAE10065 Human heat shock p
42	2013	61.5	640	21	AAE88408 Human heat shock p
43	2013	61.5	677	11	AAE09418 Hsp70 antigen from
44	2008	61.4	641	14	AAE43004 Mature mouse sperm
45	2004.5	61.2	669	11	AAE03925 T. cruzi HSP (tc70

ALIGNMENTS

RESULT 1	AAV90693 standard; Protein: 639 AA.
XX	AAV90693;
XX	29-AUG-2000 (first entry)
DE	639 residue human immunoglobulin heavy binding protein, BIP(78KD).
XX	Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;
KW	autoantigen; rheumatoid arthritis; antlerarthritis; antirheumatic; p78.
XX	
OS	Homo sapiens.
XX	
PN	WO200021995-A1.
XX	
PD	20-APR-2000.
XX	
PF	08-OCT-1999; 99WO-GB03316.
XX	
PR	09-OCT-1998; 98GB-0022115.
XX	
PA	(UNLO) KINGS COLLEGE LONDON.
PI	Paray GS, Corrigall VM, Bodman-Smith MD, Fife MS, Lanchbury JS;
XX	WPI; 2000-317942/27.
DR	N-PSDB; AAA30792.
XX	
PT	New human immunoglobulin heavy chain binding protein and encoding
PT	polynucleotide, useful for diagnosis and treatment of rheumatoid
PT	arthritis -
XX	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:18:33 ; Search time 34.99 Seconds
(without alignments)
1738.341 Million cell updates/sec

Title: US-09-806-955a-2

Perfect score: 3225

Sequence: 1 MEEDKEDVGTGVGIDLTGTT.....SKLKGSAAGPPTGEEDTAEEL 633

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	3209	99.5	654	2	A27414
2	3201	99.3	653	2	A37048
3	3200	99.2	654	1	HHRGB
4	3182.5	98.7	653	2	A29821
5	3168	98.2	652	2	T50242
6	2722.5	84.4	667	2	S24782
7	2719.5	84.3	667	2	D44261
8	2710	84.0	656	2	JN0666
9	2618.5	81.2	661	2	T15513
10	2594.5	80.4	661	2	A32475
11	2555	79.2	657	2	T34037
12	2354.5	73.0	667	2	S21879
13	2346.5	72.8	668	2	S21880
14	2332.5	72.1	663	2	T04080
15	2325	72.1	668	2	T46574
16	2322.5	72.0	663	2	T04078
17	2305.5	71.5	663	2	T03581
18	2295	71.2	668	2	S71171
19	2283	70.8	672	2	T43123
20	2276	70.6	672	2	T43123
21	2261	70.1	642	2	T45298
22	2237	69.4	655	2	H86222
23	2234.5	69.3	664	2	T06598
24	2231	69.2	658	2	S38990
25	2227.5	69.1	662	2	T50464
26	2216	68.7	663	2	T38155
27	2214.5	68.7	663	2	T05741
28	2200	68.2	666	2	T06358
29	2173	67.4	682	1	HNBK2

30	2156	66.9	663	2	S20877	dnak-type molecula
31	2108	65.4	651	2	JC7132	heat shock protein
32	2100	65.1	655	2	A48468	dnak-type molecula
33	2083	64.6	646	2	S07197	dnak-type molecula
34	2083	64.6	646	2	JC4853	dnak-type molecula
35	2083	64.6	646	2	A27077	dnak-type molecula
36	2080	64.5	641	2	PC7036	heat shock protein
37	2079	64.5	647	2	T41121	heat shock protein
38	2079	64.5	650	2	S11456	dnak-type molecula
39	2078	64.4	645	2	S41372	dnak-type molecula
40	2077	64.4	646	2	A45935	dnak-type molecula
41	2076.5	64.4	640	2	S37394	dnak-type molecula
42	2074.5	64.3	648	2	S53498	dnak-type molecula
43	2074.5	64.3	656	2	S51712	dnak-type molecula
44	2074	64.3	646	2	A35022	dnak-type molecula
45	2072	64.2	639	2	S20139	dnak-type molecula

ALIGNMENTS

RESULT 1
A27414
dnak-type molecular chaperone GRP78 precursor - Chinese hamster
N:Alternate names: glucose-regulated 78K protein; Ig heavy chain-binding protein
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999
C:Accession: A27414
R:Ring, J.; Wooden, S.R.; Kriz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.
Gene 55, 147-152, 1987
A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protei
A:Reference number: A27414, MID:87305586
A:Accession: A27414
A:Molecule type: mRNA
A:Residues: 1-654 <TIN>
C:Cross-references: GB:M17169; NID:g191090; PIDN:AA51448.1; PID:g304510
C:Genetics:
A:Gene: GRP78
C:Function:
A:Description: Involved in protein folding and assembling/disassembling of protein co
C:Superfamily: heat shock protein 70
C:Keywords: ATP; heterotetramer; immunoglobulin; molecular chaperone
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-654/Product: Ig heavy chain binding protein #status predicted <MAT>
F:651-654/Region: endoplasmic reticulum retention signal

Query Match Best Local Similarity 99.5%; Score 3209; DB 2; Length 654;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEDKEDVGTGVGIDLTGTTSCVGFNKGREVEIIANDQGNRIPSYAFTPEGERLIGDA 61
DB 20 EEDKEDVGTGVGIDLTGTTSCVGFNKGREVEIIANDQGNRIPSYAFTPEGERLIGDA 79
QY 62 AKQDLTSNPENTVPDARKLIGRTWNPDSVOODIKFLPFKYEKKTKRYIOVDIGGGGCTK 121
DB 80 AKQDLTSNPENTVPDARKLIGRTWNPDSVOODIKFLPFKYEKKTKRYIOVDIGGGGCTK 139
QY 122 FAPBEISAMVLTMMKEFAEAYLGKKVTHAVVTPAYPNDQROATKDAGTIAGLNVKRII 181
DB 140 FAPBEISAMVLTMMKEFAEAYLGKKVTHAVVTPAYPNDQROATKDAGTIAGLNVKRII 199
QY 182 NEPTAAIAVGLDREGEKNILVFEDLOGGTFDVSLLTIDNGVEVVAFTNGDTHLGEDFD 241
DB 200 NEPTAAIAVGLDREGEKNILVFEDLOGGTFDVSLLTIDNGVEVVAFTNGDTHLGEDFD 259
QY 242 QRVWEHRIKIKYKKTGGDVKKRRAVOKLRRVEKAKRAISSQHQARILESEFEGEDFS 301
DB 260 QRVWEHRIKIKYKKTGGDVKKRRAVOKLRRVEKAKRAISSQHQARILESEFEGEDFS 319
QY 302 ETTTTRAFEEIAMDLPFSTKPKVOKVLESDDLKSDIDEIVLVGGSTRIRIKIOLVKEFF 361

Db 320 ETLTRAFEEELNMDLFRSTMKPVQKVLSDSKSDIDEIVLVGSGTRIPKIQOLVKEFF 379
 Qy 362 NGKPSRGINDEAVAYGAQVAGVLSGDDPTGDLVLDVCPPLTGLTFTVGVMTKLPR 421
 Db 380 NGKPSRGINDEAVAYGAQVAGVLSGDDPTGDLVLDVCPPLTGLTFTVGVMTKLPR 439
 Qy 422 NTVPPTKKSQIFSTASDNPVTITKVEGERPLTKDNHLGLTFDLTGIPAPRGVPOIEV 481
 Db 440 NTVPPTKKSQIFSTASDNPVTITKVEGERPLTKDNHLGLTFDLTGIPAPRGVPOIEV 499
 Qy 482 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIEMVNDAEKFAEDKKLKERI 541
 Db 500 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIEMVNDAEKFAEDKKLKERI 559
 Qy 542 DTRNLESAVSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLESHODADIEDFKAK 601
 Db 560 DTRNLESAVSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLESHODADIEDFKAK 619
 Qy 602 KKELEIYQPIITSKLYGSAGPPPTGEEDTAE 632
 Db 620 KKELEIYQPIITSKLYGSAGPPPTGEEDTSE 650

RESULT 2

dnak-type molecular chaperone grp78 precursor - mouse
 N:Alternate names: B1P; glucose-regulated 78k protein; Ig heavy chain-binding protein
 C:Species: Mus musculus (house mouse)
 C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 20-Aug-1999
 C:Accession: A37048; B37048; J00094; I49137; A31934
 R:Kozakutani, Y.; Nomlington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.
 J. Cell Sci. Suppl. 11, 115-137, 1989
 A:Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulat
 A:Reference number: A37048; MUID:90130686
 A:Accession: A37048
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-655 <K21>
 A:Accession: B37048
 A:Molecule type: protein
 A:Residues: 20-41; 'X', 43-46 <K22>
 R:Perfett, C.L.J.; Hofbauer, R.; Brudynski, K.; Edwards, D.R.; Denhardt, D.T.
 Gene 82, 291-303, 1989
 A:Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero
 A:Reference number: J00094; MUID:90060818
 A:Accession: J00094
 A:Molecule type: mRNA
 A:Residues: 488-590 'G', 592-595, 'E', 597-655 <PAR>
 R:Linman, J.B.; Mote, P.L.; Walford, R.L.; Spindler, S.R.
 Gene 158, 225-229, 1995
 A:Title: Structure and regulation of the mouse GRP78 (B1P) promoter by glucose and calci
 A:Reference number: I49137; MUID:95331621
 A:Accession: I49137
 A:Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-28 <RES>
 A:Cross-references: EMBL:U06277; NID:9829364; PIDN:AAA37742.1; PID:9829365
 R:Haas, I.G.; Meo, T.
 Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988
 A:Title: cDNA cloning of the immunoglobulin heavy chain binding protein.
 A:Reference number: A31934; MUID:88176922
 A:Accession: A31934
 A:Molecule type: mRNA
 A:Residues: 514-589 'D', 591-595, 'E', 597-655 <HAA>
 A:Cross-references: GB:M19351; NID:g192231; PIDN:AAA37315.1; PID:g387113
 C:Genetics: grp78
 A:Gene: grp78
 C:Function: Involved in protein folding and assembling/disassembling of protein comp
 C:Subfamily: heat shock protein 70
 C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon
 F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-655/Product: dnak-type molecular chaperone grp78 #status experimental <MAT>
 F:652-655/Region: endoplasmic reticulum retention signal

Query Match 99.38; Score 3201; DB 2; Length 655;
 Best Local Similarity 99.48; Pred. No. 1,1e-154;
 Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EEDKKEDGVTVGIDLGTTYSQVFNKNGRVEITANDGNRTSPSYAFPEGRRLGDA 61
 Db 21 EEDKKEDGVTVGIDLGTTYSQVFNKNGRVEITANDGNRTSPSYAFPEGRRLGDA 80
 Qy 62 AKNOLTSNPENTVEDAKRLIGRTNNDPSVQODIKFLPKVEKKTKPYIYDYGCGQTKT 121
 Db 81 AKNOLTSNPENTVEDAKRLIGRTNNDPSVQODIKFLPKVEKKTKPYIYDYGCGQTKT 140
 Qy 122 FAPBEISAMVLTAKKERTAEATLGGKVTYHAYTVYAFYFADARQATKDACTAGLNWRTI 181
 Db 141 FAPBEISAMVLTAKKERTAEATLGGKVTYHAYTVYAFYFADARQATKDACTAGLNWRTI 200
 Qy 182 NEPTAAATAYGLDKREGKNTLVFDLGGTFDVSLLTIDNGVFYVATNGDTHLGEDFD 241
 Db 201 NEPTAAATAYGLDKREGKNTLVFDLGGTFDVSLLTIDNGVFYVATNGDTHLGEDFD 260
 Qy 242 QRVMEHFILKLYKKTKGDKVRKDNRAVOKLRREVEKAKRALLSSHOARIESEFYGEDFS 301
 Db 261 QRVMEHFILKLYKKTKGDKVRKDNRAVOKLRREVEKAKRALLSSHOARIESEFYGEDFS 320
 Qy 302 ETLTRAFEEELNMDLFRSTMKPVQKVLSDSKSDIDEIVLVGSGTRIPKIQOLVKEFF 361
 Db 321 ETLTRAFEEELNMDLFRSTMKPVQKVLSDSKSDIDEIVLVGSGTRIPKIQOLVKEFF 380
 Qy 362 NGKPSRGINDEAVAYGAQVAGVLSGDDPTGDLVLDVCPPLTGLTFTVGVMTKLPR 421
 Db 381 NGKPSRGINDEAVAYGAQVAGVLSGDDPTGDLVLDVCPPLTGLTFTVGVMTKLPR 440
 Qy 422 NTVPPTKKSQIFSTASDNPVTITKVEGERPLTKDNHLGLTFDLTGIPAPRGVPOIEV 481
 Db 441 NTVPPTKKSQIFSTASDNPVTITKVEGERPLTKDNHLGLTFDLTGIPAPRGVPOIEV 500
 Qy 482 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIEMVNDAEKFAEDKKLKERI 541
 Db 501 TFEIDVNGILVTAEDKGTGNKNTITNDONRLTPEIEMVNDAEKFAEDKKLKERI 560
 Qy 542 DTRNLESAVSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLESHODADIEDFKAK 601
 Db 561 DTRNLESAVSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLESHODADIEDFKAK 620
 Qy 602 KKELEIYQPIITSKLYGSAGPPPTGEEDTAE 632
 Db 621 KKELEIYQPIITSKLYGSAGPPPTGEEDTSE 651

RESULT 3

dnak-type molecular chaperone precursor - rat
 N:Alternate names: B1P; glucose-regulated 78k protein; Ig heavy chain-binding protein
 N:Contents: steroidogenesis-activator polypeptide
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-Jun-1999
 C:Accession: A23948; A60134; A26257
 R:Munro, S.; Pelham, H.R.B.
 Cell 46, 291-300, 1986
 A:Title: An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated p
 A:Reference number: A23948; MUID:66245075
 A:Accession: A23948
 A:Molecule type: mRNA
 A:Residues: 1-654 <MUN>
 A:Cross-references: GB:M14050; NID:g203150; PIDN:AAA40817.1; PID:g203151
 R:Pedersen, R.C.; Brownlie, A.C.
 Science 236, 188-190, 1987
 A:Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor.
 A:Reference number: A60134; MUID:87177981

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:25:17 ; Search time 19.71 Seconds

(without alignments)
1243.505 Million cell updates/sec

Title: US-09-806-955A-2

Sequence: 1 MEDKKEDVGTVGIDICT.....SKYGSAGPPTGEEDTAEL 633

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3216	99.7	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	99.5	654	1 GR78_MESAU	P07823 mesocricetu
3	3205	99.4	655	1 GR78_MOUSE	P20029 mus musculu
4	3200	99.2	654	1 GR78_RAT	P06761 rattus norv
5	3168	98.2	652	1 GR78_CHICK	Q90593 gallus gall
6	3069	95.2	658	1 GR78_XENLA	Q91883 xenopus lae
7	2722.5	84.4	667	1 GR78_APLCA	P29844 drosophila
8	2708	84.0	661	1 HSTC_DROME	P27420 caenorhabdi
9	2594.5	80.4	667	1 BIP4_TOBAC	Q03685 nicotiana t
10	2354.5	73.0	667	1 BIP4_TOBAC	Q03685 nicotiana t
11	2346.5	72.8	668	1 BIP3_TOBAC	Q03685 nicotiana t
12	2332.5	72.3	663	1 BIP3_MAIZE	Q24581 zea mays (m
13	2328.5	72.2	666	1 BIP_LYCES	P49118 lycopersico
14	2324.5	72.1	668	1 BIP_SPIOL	Q42424 spinacia ol
15	2322.5	72.0	663	1 BIP2_MAIZE	P24067 zea mays (m
16	2298.5	71.3	669	1 BIP1_ARATH	Q91KX3 arabidopsis
17	2296	71.2	668	1 BIP2_ARATH	Q39043 arabidopsis
18	2227.5	69.1	662	1 GR78_NEUCR	P78695 neurospora
19	2216	68.7	663	1 GR78_SCHPO	P36604 schizosacch
20	2191	67.9	670	1 GR78_YARLI	Q99170 yarowia li
21	2173	67.4	682	1 GR78_YEAST	P16474 saccharomyc
22	2100	65.1	655	1 GR78_PLAFO	Q05866 plasmodium
23	2085.5	64.7	641	1 HST1_YEAST	P10591 saccharomyc
24	2083	64.6	646	1 HSTC_HUMAN	P11142 homo sapien
25	2083	64.6	646	1 HSTC_MOUSE	P08109 mus musculu
26	2079	64.5	650	1 HSTC_BOVIN	P19120 bos taurus
27	2078	64.4	640	1 HST1_PICAN	P53421 pictia angu
28	2076.5	64.4	644	1 HSTC_DICDI	P36451 dictyosteli
29	2074.5	64.3	655	1 HST1_CANAL	P41797 candida alb
30	2074	64.3	646	1 HSTC_CANAL	P19378 citreulius
31	2072	64.2	638	1 HSTC_YEAST	P10592 saccharomyc
32	2065.5	64.0	679	1 GR78_KLULA	P22010 kluyveromyc
33	2065	64.0	648	1 HST1_PUCGR	Q01877 puccinia gr

34	2064	64.0	641	1 HST4_YEAST	P22202 saccharomyc
35	2063.5	64.0	649	1 HST3_YEAST	P09435 saccharomyc
36	2060.5	63.9	646	1 HST0_NEUCR	Q01233 neurospora
37	2060	63.9	643	1 HST1_SCHPO	Q10265 schizosacch
38	2059	63.8	643	1 HST0_SORBN	P26413 glycine max
39	2053.5	63.7	651	1 HSTC_PETRY	P09189 petunia hyb
40	2052	63.6	652	1 HST0_MANSE	Q9u639 manduca sex
41	2049	63.5	651	1 HST0_ONCMY	P08108 oncorhynch
42	2046	63.4	641	1 HST2_BOVIN	Q27965 bos taurus
43	2045	63.4	651	1 HST1_ARATH	P22953 arabidopsis
44	2043.5	63.4	639	1 HST2_HUMAN	P54652 homo sapien
45	2039.5	63.2	641	1 HST1_HUMAN	P08107 homo sapien

ALIGNMENTS

RESULT 1	ID	GR78_HUMAN	STANDARD:	PRT:	654 AA.
AC	P11021	09NPTL			
DT	01-JUL-1989	(Rel. 11, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin heavy chain binding protein) (BiP) (Endoplasmic reticulum lumenal Ca2+ binding protein grp78).				
DE	HSP45 OR GRP78.				
GN	HSP45 OR GRP78.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=8828347; PubMed=2840249;				
RA	Ting J., Lee A.S.;				
RT	"Human gene encoding the 78,000-dalton glucose-regulated protein and its pseudogene: structure, conservation, and regulation.";				
RL	DNA 7:275-286(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Cervical carcinoma;				
RA	Chao C.C.K.;				
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fibroblast;				
RA	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;				
RT	"Grp78 is involved in the quality control of the LDL-receptor.";				
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Bermudez-Fajardo A., Ulewellyn D.H., Campbell A.K., Errington R.R.;				
RT	"Sequence differences between human grp78/BiP isolated from HeLa cells and previously reported human sequences.";				
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 1-25 FROM N.A.				
RX	MEDLINE=93126087; PubMed=1480470;				
RA	Chao C.C.K., Lin-Chao S.;				
RT	"A direct-repeat sequence of the human BiP gene is required for A23187-mediated inducibility and an inducible nuclear factor binding.";				
RL	Nucleic Acids Res. 20:6481-6485(1992).				
RN	[6]				
RP	SEQUENCE OF 22-38.				
RX	TISSUE=Breast carcinoma;				
RA	MEDLINE=97295304; PubMed=9150946;				
RT	Rasmussen R.K., Ul H., Eddes J.S., Moritz R.L., Reid G.E.,				
RT	"Two-dimensional electrophoretic analysis of human breast carcinoma proteins: mapping of proteins that bind to the SH3 domain of mixed lineage kinase MLK2.";				

```

RL Electrophoresis 18:588-598(1997).
RN (7)
RP SEQUENCE OF 19-40.
RC TISSUE-Colon carcinoma;
RX MEDLINE-97295306; PubMed-9150948;
RT J. H., Reid G.E., Moritz R.L., Edges J.S., Burgess A.W., Simpson R.J.:
  "A two-dimensional gel database of human colon carcinoma proteins.";
  Electrophoresis 18:605-613(1997).
RL
-1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
  MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
EMBL: M19645; AAA52614.1; -.
DR EMBL: X87949; CAAG1201.1; -.
DR EMBL: AJ271729; CAB71335.1; -.
DR EMBL: AF216282; AAR42836.1; -.
DR EMBL: X59669; CAA42595.1; -.
DR PIR: A29821; A29821.
DR HSSP: P19120; 3HSC.
DR SWISS-2DPAGE: P11021; HUMAN.
DR PMMA-2DPAGE: P11021; -.
DR PHCI-2DPAGE: P11021; -.
DR Siena-2DPAGE: P11021; -.
DR DR MIM: 138120; -.
DR InterPro: IPR000866; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PRO0301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW ATP-binding; Endoplasmic reticulum; Signal.
FT SIGNAL 1 18
FT CHAIN 19 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT SITE 651 654 PREVENT SECRETION FROM ER.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
FT SEQUENCE 654 AA; 72333 MW; 59B7DBD85BC32A00 CRC64;
-----
Query Match 99.7%; Score 3216; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 4,6e-153;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
-----
QY 2 EEDKEDVGVGIDLTGTYSCVGVKNGRVEIIANDGNNRITPSYVAFTEGGERLIGDA 61
  |||||||
DB 20 EEDKEDVGVGIDLTGTYSCVGVKNGRVEIIANDGNNRITPSYVAFTEGGERLIGDA 79
QY 62 AKNQLTSPENTVPDAKILGRTWMDPSVODIKFLPKRYVEKKRPYIQVDIGGGQTKT 121
  |||||||
DB 80 AKNQLTSPENTVPDAKILGRTWMDPSVODIKFLPKRYVEKKRPYIQVDIGGGQTKT 139
QY 122 FAPETISAMVLTLMKETREAVLGGKKTAAVYTVPAVFNDAAQOATKDGTTAGLWVRIT 181
  |||||||
DB 140 FAPETISAMVLTLMKETREAVLGGKKTAAVYTVPAVFNDAAQOATKDGTTAGLWVRIT 199
QY 182 NEPTAAIAAYGLDKREGKRNILVFDLGGGTEDVSLITIDNGVEFVVAATNGDTHLGGEDFD 241
  |||||||
DB 200 NEPTAAIAAYGLDKREGKRNILVFDLGGGTEDVSLITIDNGVEFVVAATNGDTHLGGEDFD 259
QY 242 QRVMEHFTIKLYKKTKGDKVRKDNRAVOKLRREVEAKAKRALSSQHOARIEISFYEGEDFS 301

```

```

DB 260 QRVMEHFTIKLYKKTKGDKVRKDNRAVOKLRREVEAKAKRALSSQHOARIEISFYEGEDFS 319
  |||||||
QY 302 ETLTRAKFEELNMDLFRSGMKPVOKVLSDSLKSDIDIEIVYVGSTRIPOQLVKERF 361
  |||||||
DB 320 ETLTRAKFEELNMDLFRSGMKPVOKVLSDSLKSDIDIEIVYVGSTRIPOQLVKERF 379
QY 362 NGKERSRGINPDEAVAGAAVQAGVLSGDQDQDGLVLDVCPLTGLGIEYGVMTKLPR 421
  |||||||
DB 380 NGKERSRGINPDEAVAGAAVQAGVLSGDQDQDGLVLDVCPLTGLGIEYGVMTKLPR 439
QY 422 NTVPFTKSOIFSTASDNPVTYIVYEGEERPLTDNHLGTFDLTGIPAPRGVPOIEV 481
  |||||||
DB 440 NTVPFTKSOIFSTASDNPVTYIVYEGEERPLTDNHLGTFDLTGIPAPRGVPOIEV 499
QY 482 TFEIDVNGILRTADKDGKNNKTTITNDONRLPPEETERVNDAEKFAEDKKLKERI 541
  |||||||
DB 500 TFEIDVNGILRTADKDGKNNKTTITNDONRLPPEETERVNDAEKFAEDKKLKERI 559
QY 542 DTRNELSYAVSLKNOIGDKELGKLSSEDEKTEKAVEEKIEWLESHQDADIEDFKAK 601
  |||||||
DB 560 DTRNELSYAVSLKNOIGDKELGKLSSEDEKTEKAVEEKIEWLESHQDADIEDFKAK 619
QY 602 KKELEIYOPTISKLYGSGAPPTGEEDTAE 632
  |||||||
DB 620 KKELEIYOPTISKLYGSGAPPTGEEDTAE 650
-----
RESULT 2
GR78_MESAU STANDARD; PRT; 654 AA.
ID GR78_MESAU
AC P07823.
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin
DE heavy chain binding protein) (Bip).
GN HSP75 OR GRP78.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87305586; PubMed-3623104;
RA Ting J., Wooden S.K., Ritz R., Kelleher K., Kaufman R.J., Lee A.S.:
  "The nucleotide sequence encoding the hamster 78-kDa
  glucose-regulated protein (GRP78) and its conservation between
  hamster and rat.";
  Gene 55:147-152(1987).
RL
-1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
  MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
EMBL: M17169; AAA51448.1; -.
DR PIR: A27414; A27414.
DR HSSP: P19120; 3HSC.
DR InterPro: IPR000866; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PRO0301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00297; HSP70_1; 1.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:23:16 ; Search time 50.75 Seconds

(without alignments)
2157.750 Million cell updates/sec

Title: US-09-806-955A-2

Sequence: 1 MEEDKKEVGVVGVGIDLGTT.....SKVGSAGPPTGGEEDTAEI 633

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Minimum DB seq length: 0 562222

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:*

- 1: sp._archaea:*
- 2: sp._bacteria:*
- 3: sp._fungi:*
- 4: sp._human:*
- 5: sp._invertebrate:*
- 6: sp._mammal:*
- 7: sp._mhc:*
- 8: sp._organelle:*
- 9: sp._phage:*
- 10: sp._plant:*
- 11: sp._rodent:*
- 12: sp._virus:*
- 13: sp._vertebrate:*
- 14: sp._unclassified:*
- 15: sp._viral:*
- 16: sp._bacteriophage:*
- 17: sp._archaeo:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3225	100.0	639	4	09UK02 homo sapien
2	3199	99.2	639	11	09DC41 mus musculu
3	2994	92.8	658	13	091688 xenopus lae
4	2725	84.5	658	5	076180 bombyx mori
5	2717	84.2	656	5	09VYU3 drosophila
6	2607.5	80.9	656	5	062568 suberites d
7	2555	79.2	657	5	002085 caenorhabdi
8	2534	78.6	651	5	024798 echinococu
9	2525.5	78.3	649	5	024895 echinococu
10	2460	76.3	648	5	045038 schistosoma
11	2344.5	72.7	665	10	09M4E8 cucumis sat
12	2325.5	72.1	557	5	0961X6 trichinella
13	2325	72.1	668	10	022639 glycine max
14	2310.5	71.6	675	10	040924 pseudotsuga
15	2305.5	71.5	663	10	024182 oryza sativ
16	2291	71.0	668	10	09FSY7 corylus ave

17	2288.5	71.0	665	10	094IK4	094IK4 scherffelia
18	2283	70.8	672	3	014453	014453 aspergillus
19	2276	70.6	672	3	013280	013280 aspergillus
20	2273	70.5	668	5	09U540	09U540 toxoplasma
21	2263	70.2	669	3	09C1C1	09C1C1 cromatium
22	2261	70.1	642	5	061001	061001 toxoplasma
23	2244.5	69.6	672	3	09UNE3	09UNE3 aspergillus
24	2243	69.6	659	10	041074	041074 phaeodactyl
25	2237	69.4	655	10	004022	004022 arabidopsis
26	2234.5	69.3	664	10	039830	039830 glycine max
27	2231	69.2	658	10	012752	012752 phytophthor
28	2226.5	69.0	701	5	024928	024928 elmeria ten
29	2214.5	68.7	608	10	040058	040058 hordeum vul
30	2200	68.2	666	10	039804	039804 glycine max
31	2188.5	67.9	655	3	P87028	P87028 pneumocysti
32	2156.5	66.9	665	3	09HG01	09HG01 pichia angu
33	2149.5	66.7	669	10	094H00	094H00 oryza sativ
34	2139.5	66.3	652	3	000053	000053 pneumocysti
35	2116.5	65.6	652	5	007615	007615 plasmodium
36	2108	65.4	651	3	09UW1	09UW1 rhizopus st
37	2093	64.9	659	5	09XZ72	09XZ72 crassostrea
38	2080	64.5	642	3	09UW40	09UW40 rhizopus st
39	2079	64.5	647	3	059855	059855 schizosacch
40	2074.5	64.3	648	10	040980	040980 pisum sativ
41	2074	64.3	661	5	077164	077164 entamoeba h
42	2069	64.2	646	13	073885	073885 gallus gall
43	2063.5	64.0	647	3	094104	094104 pneumocysti
44	2062.5	64.0	645	3	094106	094106 pneumocysti
45	2057	63.8	652	5	026924	026924 trypanosoma

ALIGNMENTS

RESULT 1

ID 09UK02 PRELIMINARY; PRT; 639 AA.

AC 09UK02;

DT 01-MAY-2000 (TREMUREL.13, Created)

DT 01-MAY-2000 (TREMUREL.13, Last sequence update)

DT 01-DEC-2001 (TREMUREL.19, Last annotation update)

DE BIP PROTEIN (FRAGMENT).

OS HSPAS.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_Taxid:9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=ARTICULAR CARTILAGE;

RA File M.S.;

RT "BIP: an Autoantigen associated with Rheumatoid Arthritis.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.

DR EMBL: AF188611; AAF13605.1; -

DR HSSP: P19120; HSC.

DR InterPro: IPR001023; HSP70.

DR Pfam: PF00012; HSP70; 1.

DR PRINTS: PR00301; HEATSHOCK70.

DR PROSITE: PS00297; HSP70_1; 1.

DR PROSITE: PS00329; HSP70_2; 1.

DR PROSITE: PS01036; HSP70_3; 1.

KW ATP-binding.

FT NON_TER 639

SO SEQUENCE 639 AA; 70931 MW; 3421208PF128FD5D CRC64;

Query Match 100.0%; Score 3225; DB 4; Length 639;

Best Local Similarity 100.0%; Pred. No. 9e-155;

Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MEEDKKEVGVVGVGIDLGTTGVGVKNGRVEIANDGGRITPSYVAFTPEGERITGD 60

```
Db 1 MEEDKEDVGVGIDLGTTYSVGVFNKGRVEIANDQNRITPSVAFTEPGERLIGD 60
Oy 61 AAKNOLTSNPENTVDAKRLIGRTNDSVVOODIKFLPEKVEKTKPYIOVDIGSGOTK 120
Db 61 AAKNOLTSNPENTVDAKRLIGRTNDSVVOODIKFLPEKVEKTKPYIOVDIGSGOTK 120
Oy 121 TFAPEEISAMVLTMMKETAVALGKRVTHAVTVPAFVFNDAQROATKAGTLAGLVNRI 180
Db 121 TFAPEEISAMVLTMMKETAVALGKRVTHAVTVPAFVFNDAQROATKAGTLAGLVNRI 180
Oy 181 INEPTAAIAVGLDKREBEKNLVFDLGGGTFPVSLLTIDNCFEYVATNGDTHLGEDF 240
Db 181 INEPTAAIAVGLDKREBEKNLVFDLGGGTFPVSLLTIDNCFEYVATNGDTHLGEDF 240
Oy 241 DQVMEHEFIKLYKKTKGADVKNRAVOKLREVEKAKRALSSQOARIIESFEGEDF 300
Db 241 DQVMEHEFIKLYKKTKGADVKNRAVOKLREVEKAKRALSSQOARIIESFEGEDF 300
Oy 301 SETLTRAFKFEELNMDLFRSTMKPVQKVLSDSLKSDIDEIVLVGSGTRIPKIQOLVKEF 360
Db 301 SETLTRAFKFEELNMDLFRSTMKPVQKVLSDSLKSDIDEIVLVGSGTRIPKIQOLVKEF 360
Oy 361 FNGKPSRGINPDEAVAGAAGVAGVLSGDDOTGDLVLDVCPPLTGLIETVGVMTKLIP 420
Db 361 FNGKPSRGINPDEAVAGAAGVAGVLSGDDOTGDLVLDVCPPLTGLIETVGVMTKLIP 420
Oy 421 RNTVPTKKSQIFSTASDNOPTVTIKYVEGERPLTKDNHLGTFDLTGIPAPRGVPOLE 480
Db 421 RNTVPTKKSQIFSTASDNOPTVTIKYVEGERPLTKDNHLGTFDLTGIPAPRGVPOLE 480
Oy 481 VTPEIDVNGILRVTAEDKGTGNKKNTITTDONRLTPEIEEMVNDAEKFAEDKKLER 540
Db 481 VTPEIDVNGILRVTAEDKGTGNKKNTITTDONRLTPEIEEMVNDAEKFAEDKKLER 540
Oy 541 IDTRNELSYAALSKNOIGDKKLGKLSSEDEKTEKAVEEIKIMLESQADIEDFKA 600
Db 541 IDTRNELSYAALSKNOIGDKKLGKLSSEDEKTEKAVEEIKIMLESQADIEDFKA 600
Oy 601 KKELEIIVOPITISKLYGSAGPPPTGEDTAEI 633
Db 601 KKELEIIVOPITISKLYGSAGPPPTGEDTAEI 633

RESULT 2
O9DC41 PRELIMINARY: PRT: 655 AA.
AC O9DC41:
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
E 01-JUN-2001 (Tremblrel. 17, Last sequence update)
F 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HEAT SHOCK 70KD PROTEIN 5 (GLUCOSE-REGULATED PROTEIN, 78KD).
GN HSP45.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Aichi J., Fukuda S.,
RA Aikawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuhl P., Lewis S., Macsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brenstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guslinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
```

```
RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suuki H., Toyooka K., Mang K.H., Weitz C., Whitaker C., Wlting L.,
RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
DR EMBL: AK004578; BAB3387.1; -.
DR HSP: P19120; 3HSC.
DR MGD: MGI:95835; Hsp45.
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR01023; HSP70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW ATP-binding. 655 AA; 72422 MW; 69E2C0A2C896DE6C CRC64;
SQ
SEQUENCE 655 AA: 72422 MW; 69E2C0A2C896DE6C CRC64;

Query Match 99.2%; Score 3199; DB 11; Length 655;
Best Local Similarity 99.2%; Pred. No. 1,9e-153;
Matches 626; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 2 EEDKREDVGVGIDLGTTYSVGVFNKGRVEIANDQNRITPSVAFTEPGERLIGDA 61
Db 21 EEDKREDVGVGIDLGTTYSVGVFNKGRVEIANDQNRITPSVAFTEPGERLIGDA 80
Oy 62 AKNOLTSNPENTVPAKRLIGRTNDSVVOODIKFLPFVVKKRPYIOVDIGSGOTK 121
Db 81 AKNOLTSNPENTVPAKRLIGRTNDSVVOODIKFLPFVVKKRPYIOVDIGSGOTK 140
Oy 122 FAPBEISAMVLTMMKETAVALGKRVTHAVTVPAFVFNDAQROATKAGTLAGLVNRI 181
Db 121 FAPBEISAMVLTMMKETAVALGKRVTHAVTVPAFVFNDAQROATKAGTLAGLVNRI 200
Oy 182 NEPTAAIAVGLDKREBEKNLVFDLGGCTPVSLLTIDNCFEYVATNGDTHLGEDFD 241
Db 201 NEPTAAIAVGLDKREBEKNLVFDLGGCTPVSLLTIDNCFEYVATNGDTHLGEDFD 260
Oy 242 QVMEHEFIKLYKKTKGADVKNRAVOKLREVEKAKRALSSQOARIIESFEGEDS 301
Db 261 QVMEHEFIKLYKKTKGADVKNRAVOKLREVEKAKRALSSQOARIIESFEGEDS 320
Oy 302 ETLTRAKFEELNMDLFRSTMKPVQKVLSDSLKSDIDEIVLVGSGTRIPKIQOLVKEF 361
Db 321 ETLTRAKFEELNMDLFRSTMKPVQKVLSDSLKSDIDEIVLVGSGTRIPKIQOLVKEF 380
Oy 362 NCKEPRSRCINPDEAVAGAAGVAGVLSGDDOTGDLVLDVCPPLTGLIETVGVMTKLIP 421
Db 381 NCKEPRSRCINPDEAVAGAAGVAGVLSGDDOTGDLVLDVCPPLTGLIETVGVMTKLIP 440
Oy 422 NTVVPTKKSQIFSTASDNOPTVTIKYVEGERPLTKDNHLGTFDLTGIPAPRGVPOIEV 481
Db 441 NTVVPTKKSQIFSTASDNOPTVTIKYVEGERPLTKDNHLGTFDLTGIPAPRGVPOIEV 500
Oy 482 TPEIDVNGILRVTAEDKGTGNKKNTITTDONRLTPEIEEMVNDAEKFAEDKKLERI 541
Db 501 TPEIDVNGILRVTAEDKGTGNKKNTITTDONRLTPEIEEMVNDAEKFAEDKKLERI 560
Oy 542 DTRNELSYAALSKNOIGDKKLGKLSSEDEKTEKAVEEIKIMLESQADIEDFKA 601
Db 561 DTRNELSYAALSKNOIGDKKLGKLSSEDEKTEKAVEEIKIMLESQADIEDFKA 620
Oy 602 KKELEIIVOPITISKLYGSAGPPPTGEDTAE 632
Db 621 KKELEIIVOPITISKLYGSAGPPPTGEDTAE 651

RESULT 3
O91688 PRELIMINARY: PRT: 658 AA.
AC O91688:
ID O91688:
AC O91688:
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:17:07 ; Search time 58.8 Seconds

(without alignments)
1195.743 Million cell updates/sec

Title: US-09-806-955A-2

Perfect score: 3225
Sequence: 1 MEEDKKEVGVGVGIDLGTT.....SKLYGAGPPPGEDTAEL 633

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

1 number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

1: A.Geneseq_032802.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3225	100.0	633	21	AA90694
2	3225	100.0	639	21	AA90693
3	3209	99.5	654	15	AA83075
4	3182.5	98.7	653	22	AA812985
5	3182.5	98.3	653	22	AA82533
6	3171	98.3	656	14	AA84934
7	3168	98.2	656	15	AA853076
8	2717	84.2	656	22	AA858666
9	2296	71.2	668	21	AA81382
10	2296	71.2	691	21	AA81381
11	2258.5	70.0	642	21	AA839285

12	2258.5	70.0	658	21	AA839284	Arabidopsis thalia
13	2237	69.4	655	21	AA830912	Arabidopsis thalia
14	2173	67.4	682	22	AA870749	S cerevisiae apopt
15	2122.5	65.8	687	22	AA870875	C albicans apoptos
16	2083	64.6	646	19	AAW54354	Heat shock cognate
17	2083	64.6	646	20	AAV17408	Human heat shock c
18	2083	64.6	646	20	AAV17408	Mouse heat shock c
19	2083	64.6	646	21	AA83649	Mouse heat shock p
20	2083	64.6	646	21	AA83651	Human heat shock p
21	2083	64.6	646	22	AAE12957	Human Hsp70 family
22	2083	64.6	890	21	AA82535	Human heat shock p
23	2083	64.6	890	21	AA822938	GFP-HSC70 fusion p
24	2074.5	64.3	656	18	AAW01638	Candida albicans h
25	2064	64.0	646	11	AA803927	Rat HSP (arthritis)
26	2041	63.3	647	11	AA803928	Xenopus laevis HSP
27	2039.5	63.2	624	21	AA83252	Human Hsp72 (heat
28	2039.5	63.2	641	19	AAW54349	Human heat shock 7
29	2039.5	63.2	641	21	AA823652	Human heat shock p
30	2039.5	63.2	641	22	AAE12986	Human Hsp70 family
31	2039.5	63.2	641	22	AA82534	Human heat shock p
32	2036.5	63.1	646	21	AA83604	Arabidopsis thalia
33	2033.5	63.1	633	14	AA833002	Mouse SLIP1 homolo
34	2032.5	63.0	647	20	AAV31380	T. gondii antigen
35	2032.5	63.0	647	22	AA849099	Toxoplasma gondii
36	2032	63.0	634	11	AA803930	Gallus gallus HSP
37	2030	62.9	651	22	AA860514	Drosophila melanog
38	2021	62.7	640	11	AA833653	Human heat shock p
39	2017.5	62.6	640	21	AA803929	Homo sapiens HSP (
40	2014	62.4	642	21	AA823650	Rat heat shock pro
41	2013	62.4	640	18	AAW10065	Human heat shock p
42	2013	62.4	640	21	AAV88408	Human heat shock p
43	2013	62.4	677	11	AA809418	Hsp70 antigen from
44	2008	62.3	641	14	AA843004	Mature mouse sperm
45	2004.5	62.2	669	11	AA803925	T. cruzi HSP (tc70

ALIGNMENTS

RESULT 1	AA90694	standard; Protein: 633 AA.
ID	AA90694	
XX	AA90694:	
AC	29-AUG-2000	(first entry)
DT	633 residue human immunoglobulin heavy binding protein, BIP(78KD).	
XX		
DE	Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;	
XX	autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78.	
KW	Homo sapiens.	
XX		
OS	WO200021995-A1.	
PN	20-APR-2000.	
XX		
PD	08-OCT-1999;	99WO-GB03316.
XX		
PF	08-OCT-1998;	98GB-0022115.
XX		
PR	09-OCT-1998;	98GB-0022115.
XX		
XX	(UNLO) KINGS COLLEGE LONDON.	
PA	Panayi GS, Corrigall VM, Bodman-Smith MD, Fife MS, Lanchbury JS;	
XX	WPI: 2000-317942/27.	
PI		
XX		
DR	New human immunoglobulin heavy chain binding protein and encoding	
XX	poly nucleotide, useful for diagnosis and treatment of rheumatoid	
PT	arthritis -	
XX		
PS	Claim 3; Page 44-46; 53pp; English.	

xx This sequence represents a human immunoglobulin heavy chain binding
 cc protein, BIP(78KD), having a 633 amino acid sequence. The invention
 cc also encompasses a BIP(78KD) protein of 639 amino acids (Y90693).
 cc The cDNA encoding BIP(78KD), also referred to as p78 in the
 cc specification, was isolated from human chondrocytes (the specialised
 cc cells of articular cartilage) and human chondrosarcoma cell lines. The
 cc BIP(78KD) cDNA of this invention contains a number of differences
 cc compared with the published sequence (Genbank accession number X87949),
 cc and has therefore been deposited with Genbank with the accession
 cc number AF188611). These differences comprise 6 single nucleotide
 cc substitutions and a codon insertion, and result in three amino acid
 cc substitutions and an arginine insertion at position 834-836 of the
 cc protein. The BIP(78KD) proteins react with antibodies present in the
 cc serum of rheumatoid arthritis patients, and is therefore a putative
 cc autoantigen for this autoimmune disease. BIP(78KD) is also able to
 cc selectively proliferate synovial T-cells from patients with rheumatoid
 cc arthritis. BIP(78KD) or peptides derived from the protein are useful as
 cc reagents to indicate the presence of rheumatoid arthritis, and can be
 cc used in prognostic or diagnostic tests of body fluids for rheumatoid
 cc arthritis by ELISA (enzyme linked immunosorbent assay) or Western
 cc blotting. The protein or the cDNA encoding it can also be used to test
 cc for rheumatoid arthritis by detecting antibodies to the protein.
 cc BIP(78KD), its peptides and polynucleotides are also useful
 cc therapeutically.

xx Sequence 633 AA:

Query Match 100.0%; Score 3225; DB 21; Length 633;
 Best Local Similarity 100.0%; Pred. No. 2.8e-208; Mismatches 0; Gaps 0;
 Matches 633; Conservative 0; Indels 0;

QY 1 MEEDKEDVGTWVIGIDLTGTVSCVVFKNKGRVEITANDQGNRTSPSYAFTPEGERLIGD 60
 DB 1 meedkedvgtwvlgldltgtsvvgfkngrveitandqgnrtipsyafcegerllygd 60
 QY 61 AAKNQLTSPNPTVFDAKRLIGRTWNPDSVOQDIKFLPKVVEKKTKPYIOVDIGGGQTK 120
 DB 61 aaknqltspnptvfdaakrligrtnwpsvqdiikflpkvvekktkpyiovdigggqtk 120
 QY 121 TEAPEEISAMVLTMMKETAENALGCKVTHAVTVPAYFNDAQRQKTRKAGTLAGLWARI 180
 DB 121 tEapeeisamvltmmketaenalgckvthavtvpayfndaqrqcktagtlaglwmr1 180
 QY 181 INEPTAAIAYGLDRESEKNIPLVLDGGTFDVSLLTIDNGVFVVAATNGDTHLGGEDE 240
 DB 181 ineptaaiaYglDReSEkNIplVldGGtFDvSlltIDngVFvVAatNGdTHlgGEde 240
 QY 241 DQRVHEHTIKLYKKTKGDKVKKRRAVOKLRREVEKAKRAUSSQHOARIESFTEGEDF 300
 DB 241 dqrVhehtiklykkTKgDKvKKrRAvOKlrReVEkAKrAussqHOarIESfTEgedf 300
 QY 301 SETLIRAFEEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIVLGGSTRIRPKIQLVKEF 360
 DB 301 seTLirafEEeLNmdLfrSTmkPVokVleDSdlKsdIDeivLggSTRirPKiQLvKEf 360
 QY 361 FNGKEPSKGINPDEAVAAQAVQAGVLSGDDDTGDLVLVDCPLILGIEITVGGVWTKLIP 420
 DB 361 fngkepskginpdeavaaqavqagvlsGdddtgdlvlvdcplilgieitvGGvWtklIp 420
 QY 421 RNTVVPTRKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLGTFDPLTGIPAPRCVPOIE 480
 DB 421 rntvvpTRksQifstASdnQpTvtIKvYEGerPLtkDnHLGtFDpLTgiPAPrcvPoie 480
 QY 481 VTFEIDVNGILRVTAEDDGTGKNNKITTDQNNRLTPREIERMVNDAEKFAEDKKLER 540
 DB 481 vTFeidvngilrvtaedgtgknkkittdqnnrltpreiermvndaekfaedkkler 540
 QY 541 IDTRNELSYATSLKNOGDKKLGKLSSEDDKETMEKAVEEKIMLESODADIEDPFA 600
 DB 541 idtrnelSyatSLkNOgDKkLGkLSSEddKETMEkAVEEKIMleSODADiedPfa 600

QY 601 KKLEELIYVPIIISKLYGSAGPPPTGEDTAEL 633
 DB 601 kkleelivpIiIsKlyGSagppptgeedtael 633

RESULT 2

AAAY90693 ID AAY90693 standard; Protein: 639 AA.

AAAY90693; AC AAY90693;

DT 29-AUG-2000 (first entry)

DE 639 residue human immunoglobulin heavy binding protein, BIP(78KD).

KW Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;
 autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78.

XX Homo sapiens.

PN MO200021995-A1.

PD 20-APR-2000.

XX 08-OCT-1999; 99WO-GB03316.

PR 09-OCT-1998; 98GB-0022115.

XX (UNLO) KINGS COLLEGE LONDON.

XX Panayl GS, Corriganl VM, Bodman-Smith MD, Fife MS, Lanchbury JS;

DR WPI: 2000-317942/27.

DR N-PSDB: AAA30792.

PT New human immunoglobulin heavy chain binding protein and encoding
 PT polynucleotide, useful for diagnosis and treatment of rheumatoid
 PT arthritis.

PS Claim 3; Page 23; 53pp; English.

xx This sequence represents a human immunoglobulin heavy chain binding
 cc protein, BIP(78KD), having a 639 amino acid sequence. The invention
 cc also encompasses a BIP(78KD) protein of 633 amino acids (Y90694).
 cc The cDNA encoding BIP(78KD), also referred to as p78 in the
 cc specification, was isolated from human chondrocytes (the specialised
 cc cells of articular cartilage) and human chondrosarcoma cell lines. The
 cc BIP(78KD) cDNA of this invention contains a number of differences
 cc compared with the published sequence (Genbank accession number X87949),
 cc and has therefore been deposited with Genbank with the accession
 cc number AF188611). These differences comprise 6 single nucleotide
 cc substitutions and a codon insertion, and result in three amino acid
 cc substitutions and an arginine insertion at position 834-836 of the
 cc protein. The BIP(78KD) proteins react with antibodies present in the
 cc serum of rheumatoid arthritis patients, and is therefore a putative
 cc autoantigen for this autoimmune disease. BIP(78KD) is also able to
 cc selectively proliferate synovial T-cells from patients with rheumatoid
 cc arthritis. BIP(78KD) or peptides derived from the protein are useful as
 cc reagents to indicate the presence of rheumatoid arthritis, and can be
 cc used in prognostic or diagnostic tests of body fluids for rheumatoid
 cc arthritis by ELISA (enzyme linked immunosorbent assay) or Western
 cc blotting. The protein or the cDNA encoding it can also be used to test
 cc for rheumatoid arthritis by detecting antibodies to the protein.
 cc BIP(78KD), its peptides and polynucleotides are also useful
 cc therapeutically.

xx Sequence 639 AA:

Query Match 100.0%; Score 3225; DB 21; Length 633;
 Best Local Similarity 100.0%; Pred. No. 2.8e-208; Mismatches 0; Gaps 0;
 Matches 633; Conservative 0; Indels 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:17:45 ; Search time 26.02 seconds
(without alignments)
594,212 Million cell updates/sec

Title: US-09-806-955A-2
Perfect score: 3225
Sequence: 1 MEEDKKEVDGVTCIDLGTT.....SKLYGSAGPPPTGEEDRAEL 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 231628 seqs, 24425594 residues

Number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cg92_6/prodata/2/1aa/5A_COMB.pep:*
2: /cg92_6/prodata/2/1aa/5B_COMB.pep:*
3: /cg92_6/prodata/2/1aa/6A_COMB.pep:*
4: /cg92_6/prodata/2/1aa/6B_COMB.pep:*
5: /cg92_6/prodata/2/1aa/6CTUS_COMB.pep:*
6: /cg92_6/prodata/2/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	99.5	654	1	US-08-441-139-11 Sequence 11, Appl
2	3168	98.2	666	1	US-08-441-139-16 Sequence 16, Appl
3	2173	67.4	682	1	US-08-441-139-2 Sequence 2, Appl
4	2156	66.9	663	1	US-08-441-139-7 Sequence 7, Appl
5	2077	64.4	646	1	US-08-441-139-14 Sequence 14, Appl
6	2065	64.0	679	1	US-08-441-139-5 Sequence 5, Appl
7	1938	60.1	643	4	US-08-797-358B-3 Sequence 3, Appl
8	1547.5	48.0	679	1	US-08-214-583-2 Sequence 2, Appl
9	1474	45.7	641	1	US-08-441-139-4 Sequence 4, Appl
10	1461.5	45.3	649	4	US-09-066-047-5 Sequence 5, Appl
11	1444	44.8	607	2	US-08-472-534-5 Sequence 5, Appl
12	1303.5	40.4	600	6	5240706-1 Patent No. 5240706
13	1140.5	35.4	339	2	US-08-928-692-52 Sequence 52, Appl
14	941	29.2	187	6	5196523-13 Patent No. 5196523
15	824	25.6	168	1	US-08-441-139-10 Sequence 10, Appl
16	818.5	25.4	315	1	US-08-257-073-7 Sequence 7, Appl
17	750.5	23.3	471	1	US-08-203-905B-2 Sequence 2, Appl
18	726.5	22.5	472	4	US-08-203-905B-14 Sequence 14, Appl
19	701	21.7	307	1	US-08-858-207A-481 Patent No. 481, App
20	642.5	19.9	129	6	5196523-10 Patent No. 5196523
21	607.5	18.8	999	2	US-08-770-301A-3 Sequence 3, Appl
22	607.5	18.8	999	3	US-09-175-581-3 Sequence 3, Appl
23	598	18.5	999	2	US-08-770-301A-1 Sequence 1, Appl
24	598	18.5	999	3	US-09-175-581-1 Sequence 1, Appl
25	528.5	16.4	560	2	US-08-928-692-53 Sequence 53, Appl
26	437.5	13.6	374	2	US-08-928-692-51 Sequence 51, Appl
27	402	12.5	77	6	5196523-7 Patent No. 5196523

28	395	12.2	79	6	5196523-11 Patent No. 5196523
29	367.5	11.4	599	4	US-09-080-983-9 Sequence 9, Appl
30	358	11.1	549	2	US-08-770-544-6 Sequence 6, Appl
31	345.5	10.7	80	1	US-08-164-164-4 Sequence 4, Appl
32	345.5	10.7	80	1	US-08-338-057-4 Sequence 4, Appl
33	345.5	10.7	80	2	US-08-668-416-4 Sequence 4, Appl
34	323	10.0	554	5	PCT-US94-06430-7 Sequence 7, Appl
35	320	9.9	554	4	US-08-591-468-7 Sequence 7, Appl
36	267.5	8.3	714	2	US-08-472-534-3 Sequence 3, Appl
37	264	8.2	56	6	5196523-12 Patent No. 5196523
38	230	7.1	46	6	5196523-8 Patent No. 5196523
39	211	6.5	42	6	5196523-15 Patent No. 5196523
40	192	6.0	38	6	5196523-9 Patent No. 5196523
41	168.5	5.2	1786	4	US-08-973-462-8 Sequence 8, Appl
42	166	5.1	125	4	US-08-858-207A-407 Sequence 407, App
43	147	4.6	693	4	US-08-235-836C-68 Sequence 68, Appl
44	143.5	4.4	1805	1	US-07-853-913-2 Sequence 2, Appl
45	142.5	4.4	2285	4	US-09-308-375-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-441-139-11
; Sequence 11, Application US/08441139
; Patent No. 5773245
GENERAL INFORMATION:
; APPLICANT: Wittrup, Dr. Karl D.
; APPLICANT: Robinson, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/441,139
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEO ID NO: 11:
SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-441-139-11

Query Match 99.5%; Score 3209; DB 1; Length 654;
Best Local Similarity 99.7%; Pred. No. 3,5e-259;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Caps 0;

```

QY 2 BEDKKEDVGTWVIGIDLGTTYSVGVFKNGRVEIIANDOGNRITPSYVAFTEPGEERLIGDA 61
DB 20 BEDKKEDVGTWVIGIDLGTTYSVGVFKNGRVEIIANDOGNRITPSYVAFTEPGEERLIGDA 79
QY 62 AKNOLITSPENVTVPDAKRLIGRTWNPVSQODIKFLPKRVVEKKTKPYIYQVDSGGQTKT 121
DB 80 AKNOLITSPENVTVPDAKRLIGRTWNPVSQODIKFLPKRVVEKKTKPYIYQVDSGGQTKT 139
QY 122 FAPETISAMVLTLMKMETAEAYLGKVTAAVYVTPAYFNDAGQAKDKAGTIGLWVMRI 181
DB 140 FAPETISAMVLTLMKMETAEAYLGKVTAAVYVTPAYFNDAGQAKDKAGTIGLWVMRI 199
QY 182 NEPTAAAIAYGLDKRREGKRNILVFDLGSGTFDVSLLTIDNGVEVVAATNGDTHLGEDFD 241
DB 200 NEPTAAAIAYGLDKRREGKRNILVFDLGSGTFDVSLLTIDNGVEVVAATNGDTHLGEDFD 259
QY 242 QRVMEHFITKLYKKTKGKVRKDNRAVOKLRREVEKAKRALSQHOARIEISFEYGEDPS 301
DB 260 QRVMEHFITKLYKKTKGKVRKDNRAVOKLRREVEKAKRALSQHOARIEISFEYGEDPS 319
QY 302 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSGSTRIPKIQOLVKEFF 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSGSTRIPKIQOLVKEFF 379
QY 362 NGKEPSRGINPDEAAYGAAGVAGVLSGDQDTGDLVLDVCPLTGLGIEYVGVMTKLIPR 421
DB 380 NGKEPSRGINPDEAAYGAAGVAGVLSGDQDTGDLVLDVCPLTGLGIEYVGVMTKLIPR 439
QY 422 NTVPVTKRSQIFSTASDNOPTVTIKVYGERPLTDNHLGTFDLTGIPPARGVPOIEV 481
DB 440 NTVPVTKRSQIFSTASDNOPTVTIKVYGERPLTDNHLGTFDLTGIPPARGVPOIEV 499
QY 482 TFEIDVNGILRYTADKGTGNKNTITINDONRLTPEEIERVANDAEKFAEDDKLKERI 541
DB 500 TFEIDVNGILRYTADKGTGNKNTITINDONRLTPEEIERVANDAEKFAEDDKLKERI 559
QY 542 DTRNELESYASLKNQIGDKRGLSKLSEDEKTEMEKAVEEKIEWLESHQADIEDFKAK 601
DB 560 DTRNELESYASLKNQIGDKRGLSKLSEDEKTEMEKAVEEKIEWLESHQADIEDFKAK 619
QY 602 KKELEIVQPIITSKLYSGAGPPTEEDTAE 632
DB 620 KKELEIVQPIITSKLYSGAGPPTEEDTAE 650

```

```

RESULT 2
US-09-441-139-16
Sequence 16, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441.139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-16

```

```

Query Match 98.2%, Score 3168; DB 1: Length 666;
Best Local Similarity 97.8%, Pred. No. 9.4e-256;
Matches 617; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 BEDKKEDVGTWVIGIDLGTTYSVGVFKNGRVEIIANDOGNRITPSYVAFTEPGEERLIGDA 61
DB 32 BEDKKEDVGTWVIGIDLGTTYSVGVFKNGRVEIIANDOGNRITPSYVAFTEPGEERLIGDA 91
QY 62 AKNOLITSPENVTVPDAKRLIGRTWNPVSQODIKFLPKRVVEKKTKPYIYQVDSGGQTKT 121
DB 92 AKNOLITSPENVTVPDAKRLIGRTWNPVSQODIKFLPKRVVEKKTKPYIYQVDSGGQTKT 151
QY 122 FAPETISAMVLTLMKMETAEAYLGKVTAAVYVTPAYFNDAGQAKDKAGTIGLWVMRI 181
DB 132 FAPETISAMVLTLMKMETAEAYLGKVTAAVYVTPAYFNDAGQAKDKAGTIGLWVMRI 211
QY 182 NEPTAAAIAYGLDKRREGKRNILVFDLGSGTFDVSLLTIDNGVEVVAATNGDTHLGEDFD 241
DB 212 NEPTAAAIAYGLDKRREGKRNILVFDLGSGTFDVSLLTIDNGVEVVAATNGDTHLGEDFD 271
QY 242 QRVMEHFITKLYKKTKGKVRKDNRAVOKLRREVEKAKRALSQHOARIEISFEYGEDPS 301
DB 272 QRVMEHFITKLYKKTKGKVRKDNRAVOKLRREVEKAKRALSQHOARIEISFEYGEDPS 331
QY 302 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSGSTRIPKIQOLVKEFF 361
DB 332 ETLTRAKFEELNMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSGSTRIPKIQOLVKEFF 391
QY 362 NGKEPSRGINPDEAAYGAAGVAGVLSGDQDTGDLVLDVCPLTGLGIEYVGVMTKLIPR 421
DB 392 NGKEPSRGINPDEAAYGAAGVAGVLSGDQDTGDLVLDVCPLTGLGIEYVGVMTKLIPR 451
QY 422 NTVPVTKRSQIFSTASDNOPTVTIKVYGERPLTDNHLGTFDLTGIPPARGVPOIEV 481
DB 452 NTVPVTKRSQIFSTASDNOPTVTIKVYGERPLTDNHLGTFDLTGIPPARGVPOIEV 511
QY 482 TFEIDVNGILRYTADKGTGNKNTITINDONRLTPEEIERVANDAEKFAEDDKLKERI 541
DB 512 TFEIDVNGILRYTADKGTGNKNTITINDONRLTPEEIERVANDAEKFAEDDKLKERI 571
QY 542 DTRNELESYASLKNQIGDKRGLSKLSEDEKTEMEKAVEEKIEWLESHQADIEDFKAK 601
DB 572 DTRNELESYASLKNQIGDKRGLSKLSEDEKTEMEKAVEEKIEWLESHQADIEDFKAK 631
QY 602 KKELEIVQPIITSKLYSGAGPPTEEDTAE 632
DB 632 KKELEIVQPIITSKLYSGAGPPTEEDTAE 662

```

```

RESULT 3
US-08-441-139-2
Sequence 2, Application US/08441139
Patent No. 5773245

```